Эконометрика

Линейные модели временных рядов

Шишкин Владимир Андреевич

Пермский государственный национальный исследовательский университет

Бокс Дж., Дженкинс Г. Анализ временных рядов, прогноз и управление: Пер. с англ. // Дж. Бокс, Г. Дженкинс. — Под ред. В. Ф. Писаренко. — М.: Мир, 1974, кн. 1. — 406 с.

Модель Бокса–Дженкинса

ARMA(p, q)

Модель авторегрессии поряка p: AR(q)

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Модель скользящего среднего порядка q: MA(q)

$$y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Модель авторегрессии и скользящего среднего порядка (p, q): ARMA(p, q)

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

Модель Бокса–Дженкинса

ARMA(p, q)

$$y_t = \delta + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\theta_p(\mathsf{L}) y_t = \delta + \phi_q(\mathsf{L}) \varepsilon_t$$

Для оценивания моделей ARMA(p, q) можно использовать метод максимального правдоподобия или нелинейный метод наименьших квадратов

Модель Бокса–Дженкинса

Проверка адекватности модели ARMA(p, q)

1. Оценки коэффициентов модели должны статистически достоверно отличаться от нуля.
2. Ошибки модели являются белым шумом. Следовательно остатки регрессии e_t должны быть похожи на белый шум.

Модель Бокса–Дженкинса

Проверка адекватности модели ARMA(p, q)

1. Q-статистика Бокса–Пирса

$$Q = m \sum_{i=1}^{K} r_k^2 \sim \chi^2_{K-p-q}, \quad r_k = \frac{\sum_{t=k+1}^{n} e_t e_{t-k}}{\sum_{t=1}^{n} e_t^2}$$

Нулевая гипотеза $(\mathrm{H}_0\colon r_1 = r_2 = \dots = r_K = 0)$ отвергается, если Q больше критического значения

2. Тест Льюнга–Бокса

$$\tilde{Q} = n(n+2) \sum_{k=1}^{K} \frac{r_k^2}{n-k} \sim \chi^2_{K-p-q}$$

Модель Бокса–Дженкинса

Выбор модели ARMA(p, q)

1. Информационный критерий Акаике

$$\mathrm{AIC} = 2\frac{p+q}{n} + \ln\left(\frac{1}{n}\sum_{t=1}^{n} e_t^2\right)$$

2. Критерий Шварца

$$\frac{(p+q)\ln(n)}{n} + \ln\left(\frac{1}{n}\sum_{t=1}^{n} e_t^2\right)$$

Методология Бокса–Дженкинса

1. Идентификация модели
   1. проверка стационарности ряда
   2. построение ACF и PACF
2. Оценивание модели и проверка адекватности модели
   1. для каждой из выбранных на первом этапе моделей оцениваются их параметры и вычисляются остатки
   2. каждая модель проверяется на соответствие данным
3. Прогнозирование

Тестирование на единичные корни

$$y_t = \alpha y_{t-1} + \varepsilon_t$$

$\mathrm{H}_0 : \phi = 1$
$\mathrm{H}_1 : \phi < 1$

$$t = \frac{\hat{\phi} - \phi}{s_{\hat{\phi}}}$$

Используется расширенный тест Дики–Фуллера (ADF)

$$y_t = \alpha(y_{t-1} - y_{t-2}) + \varepsilon_t$$

Тренд и сезонность

Тренд

Долговременная тенденция изменения временного ряда.

Сезонность

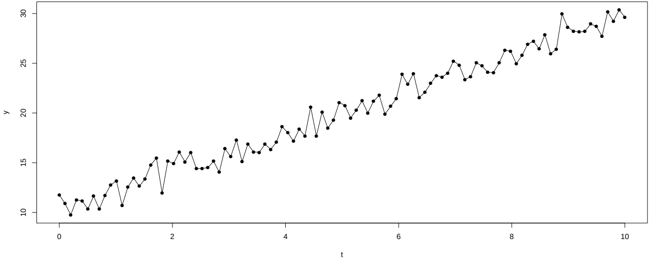
Периодическая детерминированная составляющая.

Тренд

$$y_t = f(t) + \varepsilon_t$$

$f(t)$ — тренд

ε_t — случайная составляющая

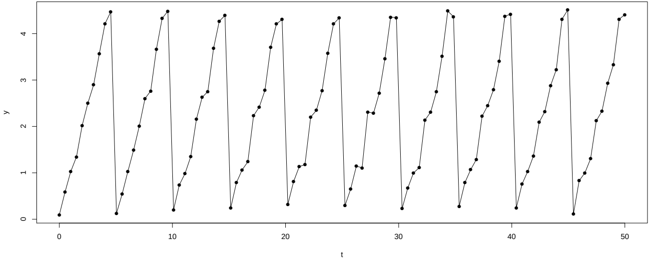


Сезонность

$$y_t = S(t) + \varepsilon_t$$

$S(t)$ — сезонная компонента: $S(t+s) \equiv S(t)$

ε_t — случайная составляющая



Разностный оператор сдвига назад

$$\mathsf{L} y_t = y_{t-1}$$

$$\nabla y_t = y_t - y_{t-1} = (1 - \mathsf{L}) y_t$$

$$\nabla^d y_t = (1 - \mathsf{L})^d y_t$$

$$\mathsf{L}^n y_t = y_{t-n}$$

$$\nabla_s^d y_t = (1 - \mathsf{L}^s)^d y_t$$

ARIMA(p, d, q)

ARMA(p, q):

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} + \varepsilon_i$$

$$\phi_p(\mathsf{L}) y_t = \theta_q(\mathsf{L}) \varepsilon_t$$

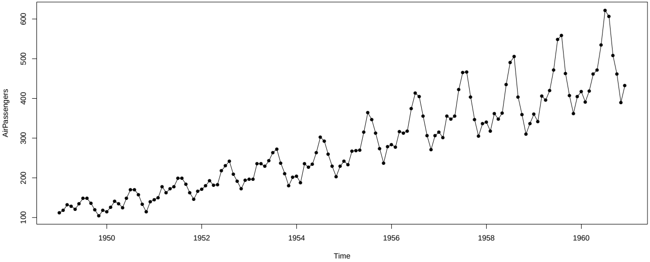
ARIMA(p, d, q):

$$\nabla^d y_t = \sum_{i=1}^{p} \alpha_i \nabla^d y_{t-i} + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} + \varepsilon_i$$

$$\phi_p(\mathsf{L}) \nabla^d y_t = \theta_q(\mathsf{L}) \varepsilon_t$$

SARIMA$(p, d, q)(P, D, Q)s$

$$\phi_p(\mathsf{L})\Phi_P(\mathsf{L}^s)\nabla^d\nabla_s^D y_t = \theta_q(\mathsf{L})\Theta_Q(\mathsf{L}^s)\varepsilon_t$$



SARIMA$(p, d, q)(P, D, Q)s$

```
arima(x,
      order = c(0, 0, 0),
      seasonal = list(order = c(0, 0, 0),
                      period = NA),
      ...)
```