Оценивание коэффициентов в моделях ARMA

Дергачёв Евгений Иванович
Крымский федеральный университет имени В.И. Вернадского
Таврическая академия
факультет математики и информатики
кафедра информатики (группа 601-И)

e-mail: evgeniydergachev@mail.ru

Для оценки коэффициентов модели авторегрессии AR(p) (не уменьшая общности, можно считать, что выборочное среднее равно 0, если это не так, надо просто вычесть его из каждого наблюдения)

$$X_t = a_1 X_{t-1} + ... + a_p X_{t-p} + \epsilon_t$$

можно применить обычный метод наименьших квадратов (МНК). Если ϵ_t - белый гауссовый шум, то значения X_t распределены нормально, а оценки коэффициентов, произведённые МНК, состоятельны и асимптотически нормальны.

$$\sum_{t=1}^{n} \epsilon_t^2 \to \min$$

$$\sum_{t=1}^{n} (X_t - a_1 X_{t-1} - ... - a_p X_{t-p})^2 \to \min$$

Для устранения неопределённости положим, что наблюдения с неположительными индексами равны 0. Воспользовавшись необходимым условием экстремума (приравняв частные производные по a_i к нулю), получим систему линейных уравнений:

$$\begin{cases} \sum_{t=1}^{n}(X_t - a_1X_{t-1} - ... - a_pX_{t-p}) * X_{t-1} & = & 0 \\ \sum_{t=1}^{n}(X_t - a_1X_{t-1} - ... - a_pX_{t-p}) * X_{t-2} & = & 0 \\ \dotfill \\ \sum_{t=1}^{n}(X_t - a_1X_{t-1} - ... - a_pX_{t-p}) * X_{t-p} & = & 0 \end{cases} \tag{0.1}$$

Решив данную систему, получим значения параметров a_i.

Попробуем применить МНК для оценки коэффициентов модели скользящего среднего MA(q)

$$X_t = b_1\epsilon_{t-1} + ... + b_q\epsilon_{t-q} + \epsilon_t$$

Снова для устранения неопределённости полагаем, что члены с неположительными индексами равны 0. Тогда

$$\epsilon_1 = X_1$$

$$\epsilon_2 = X_2 - b_1 * \epsilon_1 = X_2 - b_1 * X_1$$

$$\epsilon_3 = X_3 - b_1 * \epsilon_2 - b_2 * \epsilon_1 = X_3 - b_1 * (X_2 - b_1 * X_1) - b_2 * X_1$$

$$\epsilon_4 = X_4 - b_1 * \epsilon_3 - b_2 * \epsilon_2 - b_3 * \epsilon_1 =$$

$$= X_4 - b_1 * (X_3 - b_1 * (X_2 - b_1 * X_1) - b_2 * X_1) - b_2 * (X_2 - b_1 * X_1) - b_3 * X_1$$

Нетрудно заметить, что с каждым шагом степень полинома увеличивается на единицу, а количество слагаемых растёт экспоненциально от выбранного q, что делает данный метод оценки параметров модели неприменимым на практике в силу высокой сложности вычислений.

Для уменьшения вычислительных затрат Бокс и Дженкинс предложили использовать процедуру поиска на сетке. Идея заключается в следующем: если известны значения коэффициентов b_i, то посчитать сумму квадратов отклонений проблемы не составляет. Значит можно перебрать значения коэффициентов и выбрать лучшие в плане минимизации суммы квадратов отклонений.

Для оценки параметров ARMA(p, q) может применяться комбинация метода наименьших квадратов с поиском на сетке.

Введём оператор сдвига L:

$$LX_t = X_{t-1}$$

Тогда модель ARMA(p, q) можно записать в следующем виде:

$$(1 - a_1L - ... - a_pL^p)X_t = (1 + b_1L + ... + b_qL^q)\epsilon_t$$

или

$$X_t = \frac{(1 + b_1L + ... + b_qL^q)\epsilon_t}{1 - a_1L - ... - a_pL^p}$$

Введём вспомогательный случайный процесс:

$$Z_t = \frac{\epsilon_t}{1 - a_1L - ... - a_pL^p}$$

Тогда

$$X_t = (1 + b_1L + ... + b_qL^q)Z_t$$

Выразим отсюда Z_t, полагая члены с неположительными индексами равными нулю.

$$Z_1 = X_1$$
$$Z_2 = X_2 - b_1Z_1$$
$$Z_3 = X_3 - b_1Z_2 - b_2Z_1$$

....

Найдём значения параметров b_i с помощью процедуры поиска на сетке.

$$Z_t = a_1Z_{t-1} + a_2Z_{t-2} + ...a_pZ_{t-p} + \epsilon_t$$

Относительно процесса Z_t модель стала авторегрессионной, и параметры a_i можно оценить с помощью МНК.

Список литературы

[1] Канторович Г.Г. *Экономический журнал Высшей школы экономики. Том 6, №2.* – Москва, – 2002.