Авторегрессионные ряды: поиск единичных корней и проверка на коинтегрированность.

Власов Сергей Владимирович, 522-я группа

Санкт-Петербургский Государственный Университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель — к.ф.-м.н. **Т.М. Товстик**
Рецензент — к.ф.-м.н. **А.Ф. Сизова**

Санкт-Петербург
2007г.

- $AR(p) : X_t = \alpha + \beta t + a_1 X_{t-1} + a_2 X_{t-2} \ldots + a_p X_{t-p} + \varepsilon_t,\ t = 1, \ldots n$ — авторегрессия порядка p с линейным трендом, $\varepsilon \sim N(0, \sigma^2)$.

Теорема

Процесс x_t стационарен, если корни характеристического полинома $\chi(y)$ находятся вне единичного круга:

$$\chi(y) = \sum_{k=0}^{N} a_k y^k, \quad a_0 = 1$$

- Временной ряд X_t называется **интегрированным** $I(k)$ **порядка** k, если:
  1. ряд X_t не является стационарным в широком смысле;
  2. ряд $\triangle^k X_t$, полученный в результате k-кратного дифференцирования ряда X_t, является стационарным рядом;
  3. ряд $\triangle^{k-1} X_t$, полученный в результате $(k-1)$-кратного дифференцирования ряда X_t, не является стационарным в широком смысле.

Основные определения и обозначения

Замечание

Временной ряд является интегрируемым рядом степени k, если характеристический многочлен имеет k единичных корней, а остальные корни находятся вне единичного круга.

- Если характеристический многочлен имеет корень внутри единичного круга отличный от 1, то процесс имеет «взрывной» характер.
- Эффект коинтеграции впервые описан Грейнджером, 1981г. :

Определение

*Ряды X_t и Y_t $(X_t, Y_t \in I(1))$ называются **коинтегрированными**, если существует ненулевой **(коинтегрирующий)** вектор $\beta = (\beta_1, \beta_2)^T \neq 0$, для которого $\beta_1 X_t + \beta_2 Y_t$ — стационарный ряд.*

Постановка задачи

- Предпосылки:
  - даны 2 выборки, состоящие из реальных данных;
  - построение адекватных авторегрессионных моделей для обоих выборок;
  - выявление наличия и количества единичных корней в каждой модели;
  - проверка наличия коинтеграционного эффекта между ними.
- Задача:

  построить единый многофакторный алгоритм проверки данных на единичный корень и коинтеграционный эффект.
- Свойства алгоритма:
  - упрощает существующие методы,
  - позволяет проверять вышеописанные эффекты для любых данных.

Результаты. Содержание работы

Структура алгоритма:

- Первоначальная оценка степени авторегрессии, коэффициентов и наличия единичных корней методом МНК.
- Проверка гипотезы о наличии единичного корня в соответствии с первоначальной оценкой с использованием алгоритмов Дикки-Фуллера, Дикки-Пантула, Доладо и др.
- Проверка двух временных рядов на коинтеграцию в случае принадлежности к классу I(1).

Этап №1. МНК

- Проведем последовательные оценки коэффициентов модели методом МНК в предположении, что порядок авторегрессии равен $1, 2, \ldots, p$.

- Пусть:
$\hat{\alpha}^{(1)}, \hat{\beta}^{(1)}, \hat{a_1}^{(1)}$ — оценки параметров AR(1);
$\hat{\alpha}^{(2)}, \hat{\beta}^{(2)}, \hat{a_1}^{(2)}, \hat{a_2}^{(2)}$ — оценки параметров AR(2);
...
$\hat{\alpha}^{(p)}, \hat{\beta}^{(p)}, \hat{a_1}^{(p)}, \hat{a_2}^{(2)} \ldots \hat{a_p}^{(p)}$ — оценки параметров AR(p).

- Определение порядка авторегрессии с использованием выборочной автокорреляционной функции $(r_{part}(p) = \hat{a_p}^{(p)})$.

Утверждение

Если X_t — процесс типа AR(p), то при больших и $k > p$ распределение $r_{part}(k)$ можно аппроксимировать нормальным распределением $r_{part}(k) \sim N(0, T^{-1})$.

- **Предварительной** оценкой порядка авторегрессии является такое p, что:

$$|\hat{a_p}^{(p)}| > \frac{2}{\sqrt{N}}, |\hat{a}_{p+1}^{(p+1)}| < \frac{2}{\sqrt{N}}.$$

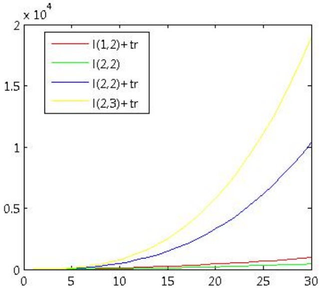
Этап №1. МНК.

Наблюдение

При решении характеристического уравнения с полученными с помощью МНК коэффициентами, получаем правильное количество единичных корней независимо от степени авторегрессии, наличия тренда и точности оценок коэффициентов.

Таблица: Результаты использования МНК:

p	k	тренд	оценка коэф	оценка p	кол-во единичных корней
3	0	-	+	+	+
3	0	+	+	+	+
1	1	-	+	+	+
1	1	+	+	+	+
2	1	-	+	+	+
2	1	+	+	+	+
3	1	-	+	+	+
3	1	+	+	+	+
2	2	-	+	+	+
2	2	+	+	-	+
3	2	-	+	-	+
3	2	+	+	-	+
4	2	-	-	-	+
4	2	+	-	-	+



Пусть r - число единичных корней.

1. При $r = 0$ или $r = 1$, $p = 1$. Следует использовать многовариантную процедуру Доладо проверки единичного корня.
2. При $r = 0$ или $r = 1$, $p > 1$. Следует использовать расширенный критерий Дикки-Фуллера проверки наличия единичного корня.
3. При $r = 2$, $p = 2$. Следует использовать процедуру Дикки-Пантулы;
4. При $r = 2$, $p > 2$ и очень сильном росте, можно с уверенностью говорить о соответствии данных модели I(2). При не очень сильном росте есть вероятность, что это модель I(1) + тренд. Чтобы проверить это, берем первые разности и проверяем на стационарность.

Многовариантная процедура Доладо проверки наличия единичного корня состоит из последовательного перебора различных комбинаций оцениваемой статистической модели data **SM** (statistical model) и процесса порождения данных **DGP** (data generating process). Для каждого шага строится статистика проверки соответствующей гипотезы по SM. Сравнение с критическим значением статистики по DGP приводит к одному из двух вариантов, либо мы отвергаем гипотезу о том, что $a_1 = 1$, либо переходим к следующему шагу проверки.

Шаги процедуры Доладо:

1. $SM:\ \Delta X_t = \alpha + \beta t + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \alpha + \varepsilon_t;\ H_0 : a_1 = 0.$
2. $SM:\ \Delta X_t = \alpha + \beta t + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \alpha + \beta t + \varepsilon_t, \beta \neq 0;\ H_0 : \beta = 0.$
   $SM:\ \Delta X_t = \alpha + \beta t + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \alpha + \beta t + \varepsilon_t,\ \beta \neq 0;\ H_0 : a_1 = 0.$
3. $SM:\ \Delta X_t = \alpha + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \varepsilon_t;\ H_0 : a_1 = 0.$
4. $SM:\ \Delta X_t = \alpha + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \varepsilon_t;\ H_0 : \alpha = 0.$
   $SM:\ \Delta X_t = \alpha + a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \alpha + \varepsilon_t,\ \alpha \neq 0;\ H_0 : a_1 = 0.$
5. $SM:\ \Delta X_t = a_1 X_{t-1} + \varepsilon_t;\ DGP:\ \Delta X_t = \varepsilon_t;\ H_0 : a_1 = 0.$

- Процедура Доладо подходит исключительно для моделей AR(1). Расширенный критерий Дикки-Фуллера позволяет проверить наличие единичного корня для любого p.

- Стандартное уравнение авторегрессии

$$X_t = \alpha + \beta t + a_1 X_{t-1} + a_2 X_{t-2} \dots a_p X_{t-p} + \varepsilon_t$$

приводим к следующему виду:

$$X_t = \alpha + \beta t + \rho X_{t-1} + (\theta_1 \Delta X_{t-1} + \ldots + \theta_{p-1} \Delta X_{t-p+1}) + \varepsilon_t,$$

где

$$\rho = a_1 + a_2 + \ldots + a_p,\ \theta_j = -(a_{j+1} + \ldots + a_p).$$

- Гипотеза о наличии единственного единичного корня принимает вид $H_0 : \rho = 1$ (Hamilton, 1994).

Этап №2. Процедура Дикки-Пантула

- Процедура Дики-Пантула состоит в проверке гипотезы, что все p корней характеристического многочлена единичные, при ее отвержении — $p-1$ корней и т.д.

- Пример использования для AR(2):
  - $AR(2) : \Delta^2 X_t = (a-1)(1-b)X_{t-1} + (ab-1)\Delta X_{t-1} + \varepsilon_t.$
    где $a = 1/z_1,\ b = 1/z_2,$ а z_1, z_2 — корни характеристического уравнения $\chi(z) = 0;$
  - Если $a = b = 1$, то $\Delta^2 X_t = \varepsilon_t.$

Гипотеза об I(2) имеет вид $H_0 : \varphi = 0$ в модели $\Delta^2 X_t = \alpha + \varphi \Delta X_{t-1} + \varepsilon_t.$

Этап №3. Проверка на коинтегрированность. Метод Энгла-Грейнджера

- Метод Энгла-Грейнджера проверяет остатки на стационарность.

- Алгоритм метода:

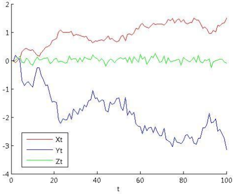
  1. Оцениваем значения α и β в рамках регрессионной модели $Y_t = \alpha + \beta X_t + u_t$. Получаем оценки для остатков: $Z_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$.

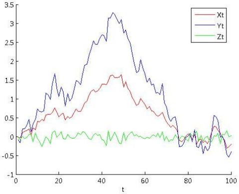
  2. Гипотеза о некоинтегрированности X_t и Y_t соответствует гипотезе о наличии единичного корня у Z_t. Проверяем используя расширенный критерий Дикки-Фуллера:

$$\Delta Z_t = \varphi Z_{t-1} + \theta_1 \Delta Z_{t-1} + \zeta_t; H_0 : \varphi = 1$$

- Статистика Энгла-Грейнджера представляет собой обычную t-статистику для проверки гипотезы $\varphi = 1$. Распределение статистики Энгла-Грейнджера отличается от стандартной статистики Дикки-Фуллера.

- Рассмотрим следующие временные ряды:
  (1): $X_t = X_{t-1} + \varepsilon_t,\ Y_t = 2X_t + \nu_t,\ t = 1,\dots,N; N = 100.$
  (2): $X_t = X_{t-1} + \varepsilon_t,\ Y_t = -2X_t + \nu_t,\ t = 1,\dots,N; N = 100.$





- Результаты МНК:

	AR(1)	AR(2)	AR(3)	r
X_t	(-0.01,0,0.96)	(-0.01,0,1.02,-0.06)	(-0.01,0,1.02,-0.01,-0.05)	1
Y_t	(-0.09,0,0.98)	(-0.09,0,1.07,-0.1)	(-0.09,0,1.08,-0.12, 0.01)	1

- Предварительный вывод: оба временных ряда являются AR(1), I(1).
- Проведение процедуры Доладо подтверждает этот факт.
- Оценка регрессии:(1): $Y_t = -0.006 + 1.983X_t + \hat{u}_t$, (2): $Y_t = -2X_t + \hat{u}_t$
  (1): $Z_t = \hat{u}_t = Y_t + 0.006 - 1.983X_t$, (2): $Z_t = \hat{u}_t = Y_t + 2X_t$
- Значение статистики Энгла-Грейнджера
  (1): $t_\varphi = -7.42 < t_{cr,0.05} = -3.396$; (2): $t_\varphi = -6.93 < t_{cr,0.05}$.
- Отвергаем гипотезу о некоинтегрированности рядов X_t и Y_t.

Заключение

Результаты работы:

- найден новый метод нахождения количества единичных корней в моделях авторегрессии.
- построены критические значения статистик для проверки гипотез о наличии единичных корней.
- построен единый многофакторный алгоритм поиска единичных корней и проверки на коинтеграцию.