Проверка корреляционным методом наличия единичного корня у процесса авторегрессии

Буйнова Светлана Николаевна, гр. 522

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н., доц. Т.М.Товстик
Рецензент: к.ф.-м.н., доц. А.Ф.Сизова

Санкт-Петербург
2010.

Основные определения

Процесс авторегрессии p-го порядка (AR(p)) задается отношением:

$$Y_t = a_1 Y_{t-1} + \ldots + a_p Y_{t-p} + \varepsilon_t,\ t = 1, 2, \ldots,$$

Процесс авторегрессии p-го порядка с линейным трендом задается отношением:

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

где ε_t - процесс белого шума с $D(\varepsilon_t) = \sigma_\varepsilon^2$, $a_p \neq 0$.

Теорема

Процесс Y_t стационарен, если все корни характеристического многочлена

$$Q(y) = 1 - a_1 y - \ldots - a_p y^p = -a_p (y - y_1) \ldots (y - y_p)$$

по модулю больше 1.

Если характеристический многочлен имеет корень, который по модулю меньше единицы, то процесс имеет „взрывной“ характер.

Основные определения

Определение

Временной ряд X_t называется ***стационарным относительно детерминированного тренда*** *$f(t)$, если $X_t - f(t)$ стационарный.*

Если ряд X_t стационарен относительно некоторого детерминированного тренда, то говорят, что этот ряд принадлежит классу рядов, стационарных относительно детерминированного тренда, или что он является TS рядом (TS — time stationary). В класс TS рядов включаются также стационарные ряды, не имеющие детерминированного тренда.

Основные определения

Под дифференцированием ряда будем понимать:

$$\Delta^k X_t = \underbrace{\Delta(\Delta(\ldots(\Delta X_t)))}_{k};\ \Delta X_t = X_t - X_{t-1}.$$

Определение

Временной ряд X_t называется ***интегрированным порядка*** *k (k = 1, 2, …) или $I(k)$, если:*

- *ряд X_t не является TS рядом,*
- *ряд $\Delta^k X_t$, полученный в результате k-кратного дифференцирования является стационарным рядом,*
- *ряд $\Delta^{k-1} X_t$, полученный в результате $(k-1)$-кратного дифференцирования не является TS рядом.*

Совокупность интегрированных рядов различных порядков $k = 1, 2, \ldots$ образуют класс разностно стационарных (DS — difference stationary) рядов.

Постановка задачи

Дано: выборка, состоящая из реальных данных.
Задачи:

- подобрать адекватную авторегрессионную модель, описывающую поведение наблюдаемого ряда,
- выявить наличие единичного корня в модели корреляционным методом и с помощью критерия Дики–Фуллера,
- сравнить результаты, полученные двумя методами.

Структура алгоритма определения наличия единичного корня:

1. оценка порядка авторегрессии, коэффициентов модели с помощью МНК.
2. выделение тренда у ряда.
3. проверка гипотезы о наличии единичного корня по критерию Дики–Фуллера, корреляционным методом.

Процесс авторегрессии первого порядка

Рассматриваемые модели:

- процесс

$$X_t^1 = \alpha + \beta t + a_1 X_{t-1}^1 + \varepsilon_t,\ |a_1| < 1 \qquad (1)$$

принадлежит классу TS рядов, стационарный относительно детерминированного тренда

$$\mu + \gamma t = \frac{\alpha - a_1(\alpha + \beta)}{(1 - a_1)^2} + \frac{\beta}{1 - a_1} t.$$

- процесс

$$X_t^2 = \beta + X_{t-1}^2 + \varepsilon_t \qquad (2)$$

принадлежит классу DS рядов.

Проверка наличия единичного корня для авторегрессии первого порядка

Критерий Дики–Фуллера.

$$H_0 : a_1 = 1,\ H_A : a_1 < 1;\ t_{stat} = \frac{\hat{a_1} - 1}{s(\hat{a_1})}.$$

- $t_{stat} < t_{crit} \Rightarrow$ гипотеза о наличии единичного корня отвергается,
- $t_{stat} > t_{crit} \Rightarrow$ гипотеза о наличии единичного корня не отвергается.

Корреляционный метод.

Корреляции процесса $Y_t^1 = X_t^1 - \mu - \gamma t = a_1 Y_{t-1}^1 + \varepsilon_t,\ |a_1| < 1$:

$$Corr(k) = a_1^k.$$

Корреляции процесса $Y_t^2 = X_t^2 - \beta t = Y_{t-1}^2 + \varepsilon_t$:

$$Corr(t,k) = \sqrt{\frac{t}{t+k}}.$$

Процесс авторегрессии второго порядка

Рассматриваемые модели:

- процесс

$$X_t^1 = \alpha + \beta t + a_1 X_{t-1}^1 + a_2 X_{t-2}^1 + \varepsilon_t,\ |x_i| > 1,\ i = 1, 2, \quad (3)$$

  (где x_i — корни характеристического многочлена $Q(x) = 1 - a_1 x - a_2 x^2$) принадлежит классу TS рядов, стационарный относительно детерминированного тренда

$$\mu + \gamma t = \frac{\alpha - a_1(\alpha + \beta) - a_2(\alpha + 2\beta)}{(1 - a_1 - a_2)^2} + \frac{\beta}{1 - a_1 - a_2} t.$$

- процесс

$$X_t^2 = \beta + a_1 X_{t-1}^2 + a_2 X_{t-2}^2 + \varepsilon_t,\ |x_1| = 1,\ |x_2| > 1, \quad (4)$$

  (где x_i — корни характеристического многочлена $Q(x) = 1 - a_1 x - a_2 x^2$) принадлежит классу DS рядов.

Проверка наличия единичного корня для авторегрессии второго порядка

Расширенный критерий Дики–Фуллера.

$$(3) \Leftrightarrow X_t^1 = \alpha + \beta t + cX_{t-1}^1 + d\Delta X_{t-1}^1 + \varepsilon_t,\ c = a_1 + a_2,\ d = -a_2. \qquad (5)$$

$$H_0 : c = 1, H_A : c < 1;\ t_{stat} = \frac{\hat{c}-1}{s(\hat{c})}.$$

- $t_{stat} < t_{crit} \Rightarrow$ гипотеза о наличии единичного корня отвергается,
- $t_{stat} > t_{crit} \Rightarrow$ гипотеза о наличии единичного корня не отвергается.

Корреляционный метод.

Корреляции процесса $Y_t^1 = X_t^1 - \mu - \gamma t = a_1 Y_{t-1}^1 + a_2 Y_{t-1}^1 + \varepsilon_t$, представляют либо экспоненту, либо затухающую синусоиду.

Корреляции процесса $Y_t^2 = X_t^2 - \beta t = Y_{t-1}^2 + d\Delta Y_{t-2}^2 + \varepsilon_t$:

$$Corr(t,k) = \frac{\left(t + \frac{2d^{t+1} - 2d^2 - d^k + d + d^{t+k+1} - d^{t+2}}{(d-1)^2} + \frac{d^{t+1} - d^k - 2d(t-1)}{d-1}\right)}{\sqrt{t + 2\left(\frac{d^2(d^{t-1}-1)}{(d-1)^2} - \frac{(t-1)d}{d-1}\right)}\sqrt{t+k+2\left(\frac{d^2(d^{t+k-1}-1)}{(d-1)^2} - \frac{(t+k-1)d}{d-1}\right)}}.$$

Процесс авторегрессии второго порядка класса TS

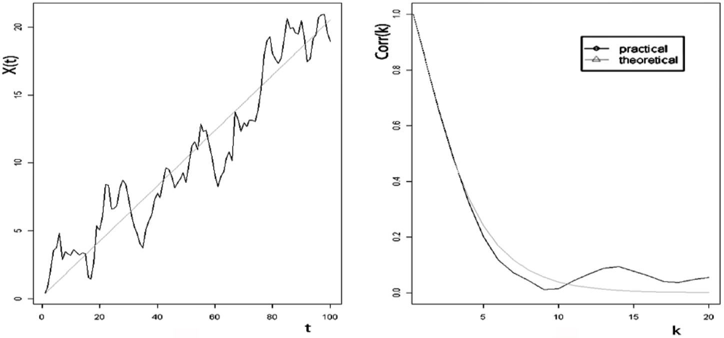

Рис.: $X_t = 0.2 + 0.04t + 1.16X_{t-1} - 0.33X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0,1)$.

Моделировался ряд: $X_t = 0.2 + 0.04t + 1.16X_{t-1} - 0.33X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0,1)$.
Подобрана модель: $X_t = 0.14 + 0.038t + 1.08X_{t-1} - 0.27X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0,0.95)$.
Корни характеристического уранения: $x_1 = 1.45$, $x_2 = 2.54$.
$t_{stat} = -1.03 > -3.45 = t_{crit} \Rightarrow$ по критерию Дики–Фуллера гипотеза H_0 не отвергается.
Оцененные корреляции ряда без тренда убывают как корреляции стационарного процесса.

Процесс авторегрессии второго порядка класса DS

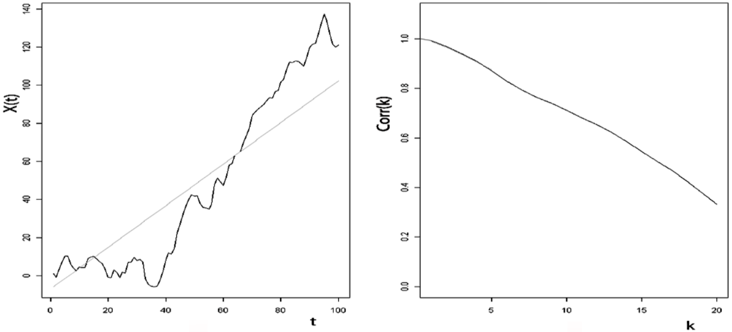

Рис.: $X_t = 0.2 + 1.5X_{t-1} - 0.5X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0, 9)$.

Моделировался ряд:
$X_t = 0.2 + 1.5X_{t-1} - 0.5X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0, 9)$.
Подобрана модель: $X_t = 0.27 + 1.454X_{t-1} - 0.472X_{t-s} + \varepsilon_t$, $\varepsilon_t \sim N(0, 8.31)$.
Корни характеристического уравнения: $x_1 = 1.04$, $x_2 = 2.04$.
$t_{stat} = -0.13 > -3.45 = t_{crit} \Rightarrow$ по критерию Дики–Фуллера гипотеза о наличии единичного корня не отвергается.
Оцененные корреляции ряда без тренда убывают очень медленно, из чего можно сделать вывод, что уравнение имеет единичный корень.

Результаты

- моделировались процессы авторегрессии X_t первого и второго порядка с линейным трендом при различных значениях N и σ_ε^2
- считалась t_{stat} и делался вывод о наличии единичного корня по критерию Дики–Фуллера.
- корреляционным методом проводился анализ и делался вывод о существовании единичного корня в модели.

В таблицах ниже представлено количество правильно сделанных выводов (из 100) для AR(1) и AR(2):

$N = 100$	$\sigma_\varepsilon^2 = 1$		$\sigma_\varepsilon^2 = 4$	
a_1	Корр.	Д.–Ф.	Корр.	Д.–Ф.
0.5	100	70	100	9
0.6	100	40	100	6
0.7	100	21	100	5
0.8	100	15	100	2
0.85	98	10	94	1
0.9	86	4	84	1
0.92	65	3	78	1
0.95	18	1	24	1
1	75	100	77	100
$N = 500$	$\sigma_\varepsilon^2 = 1$		$\sigma_\varepsilon^2 = 4$	
a_1	Корр.	Д.–Ф.	Корр.	Д.–Ф.
0.85	100	100	100	97
0.9	100	95	87	66
0.92	86	62	61	38
1	100	100	100	100

Таблица: AR(1).

N	Корр.	Д.–Ф.
100	100	0
250	100	0
300	100	5
350	100	8
400	100	22
450	100	46

Таблица: AR(2), $a_1 = 1.16$, $a_2 = -0.33$, $\sigma^2_{\varepsilon_t} = 1$.

N	Корр.	Д.–Ф.
100	78	100
250	97	100
300	98	100
350	100	100

Таблица: AR(2), $a_1 = 1.5$, $a_2 = -0.5$, $\sigma^2_{\varepsilon_t} = 1$.

Реальные данные

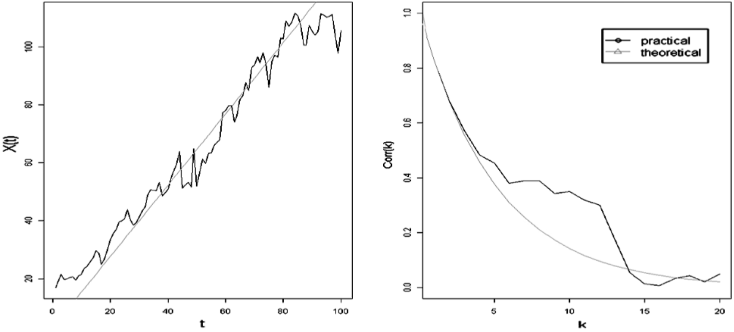

Рис.: Данные по объему услуг связи в России с октября 2001 по март 2010 в млрд. руб.

Приведены реальные данные. Подобрана модель:

$$X_t = 1.55 + 0.21t + 0.82X_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, 15.43).$$

Корень характеристического уравнения: $x_1 = 1.21$.
$t_{stat} = -2.38 > -3.45 = t_{crit} \Rightarrow$ по критерию Дики–Фуллера гипотеза о наличии единичного корня не отвергается.
Оцененные корреляции ряда без тренда убывают как корреляции стационарного процесса, значит уравнение не имеет единичного корня.

Заключение

Результаты работы:

- Показано, что наряду с проверками гипотез о наличии единичного корня у характеристического уравнения процесса авторегрессии можно использовать корреляционный метод для принятия решения, к какому классу рядов (TS или DS) отнести исследуемый ряд.
- Исходя из всего сказанного выше, при довольно большой дисперсии ошибки и маленьком количестве наблюдений целесообразнее использовать корреляционный метод.