Тема 5.

Регрессионный анализ со стационарными ВР

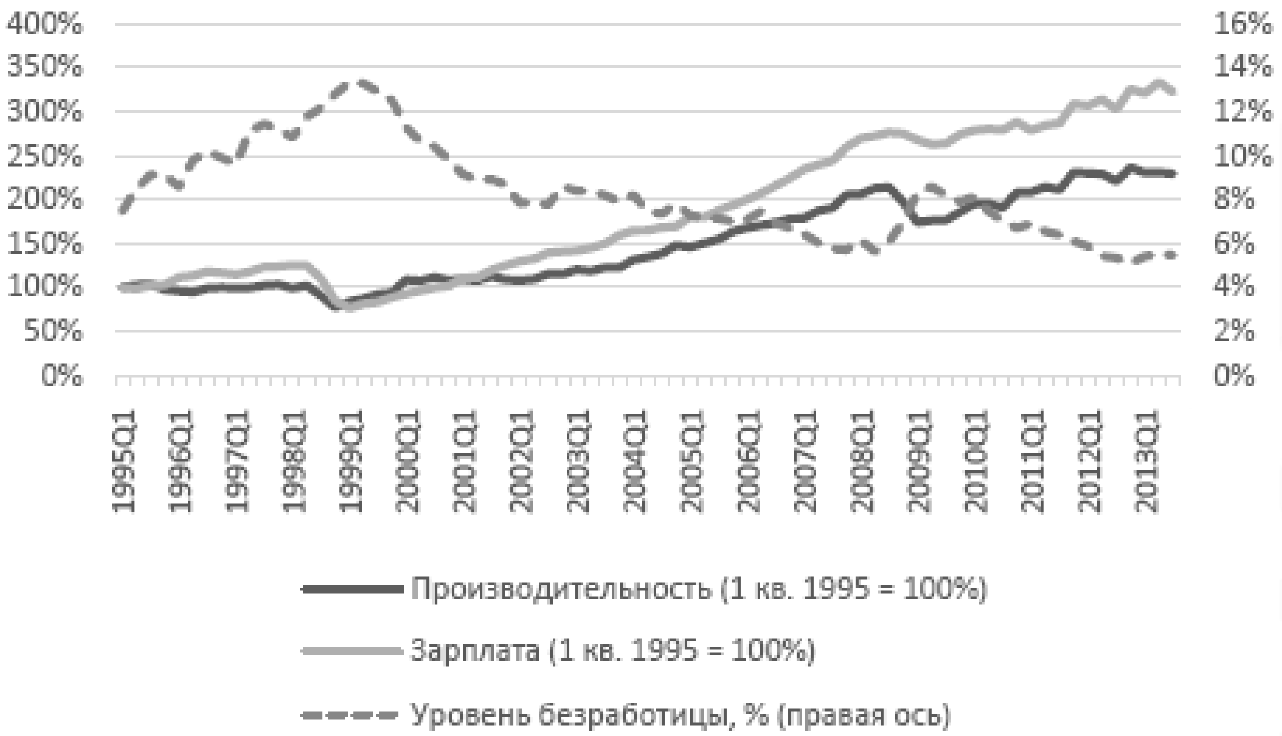
Можно ли рассматривать сразу несколько временных рядов?

Примеры

1. Резкие изменения цен на нефть могут помочь в объяснении потребления бензина. ***Прогнозирование:*** какой уровень потребления бензина ожидается в будущем, если в предыдущие два года цены на нефть снизятся более чем на 10%?

2. Как доходы населения влияют на потребление определенных товаров? И т.д.

Примеры



Динамика производительности труда, средней заработной платы и уровня безработицы, сглаженные на сезонность

Вакуленко 2014

Примеры

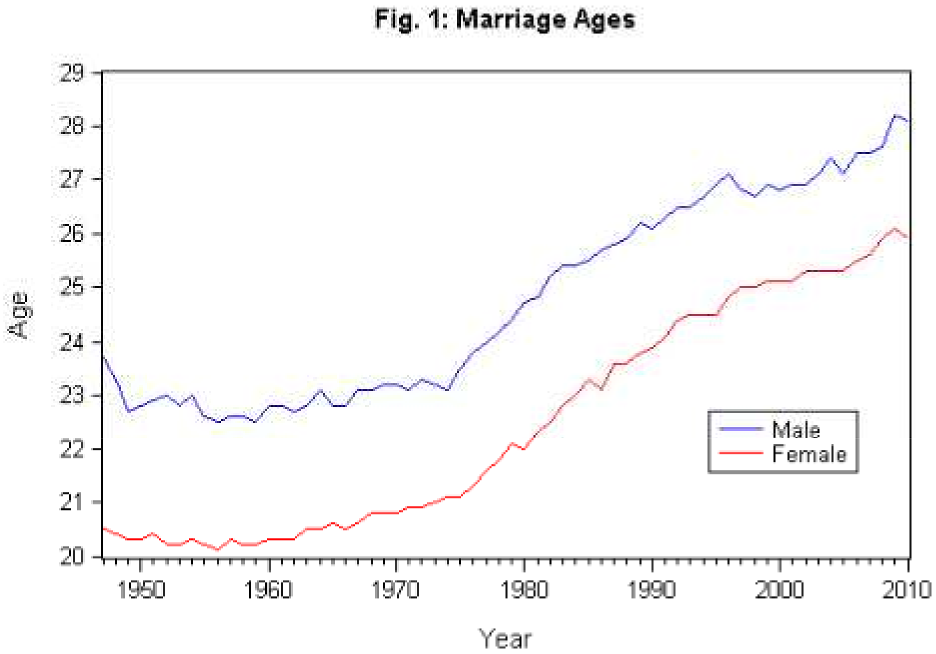
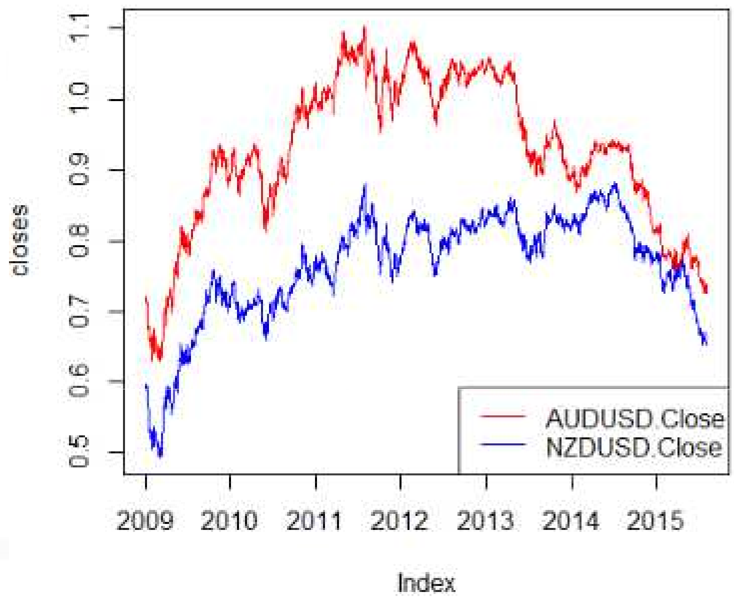
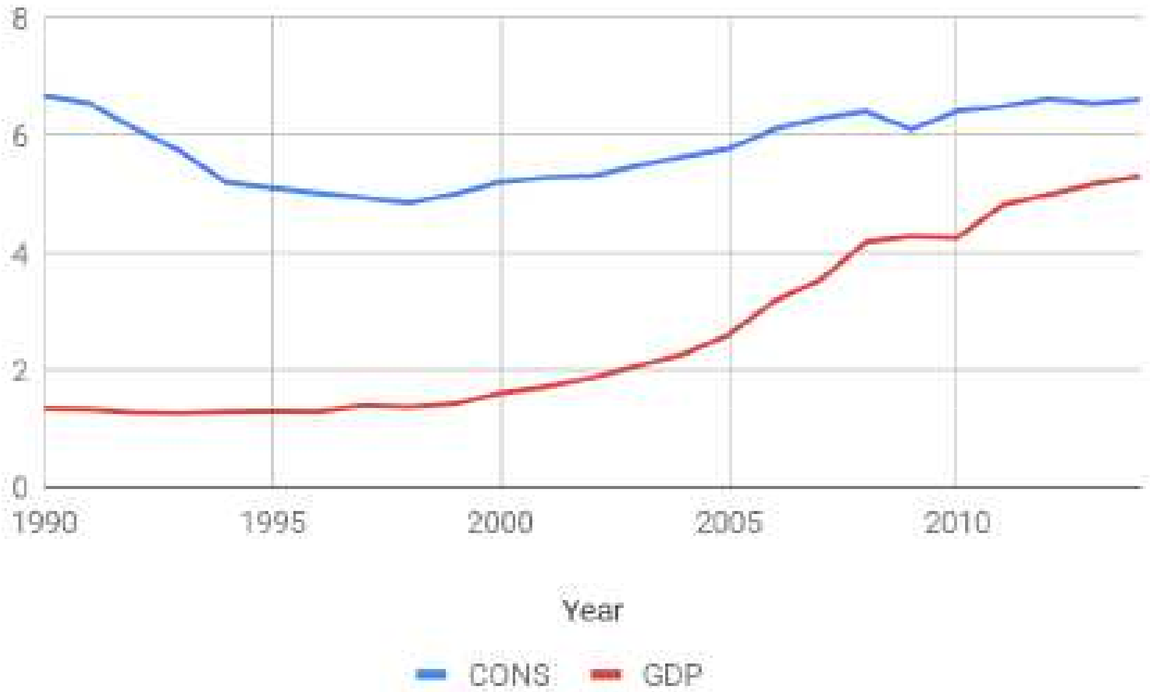


Fig. 1: Marriage Ages
Age
Year
Male
Female
20
21
22
23
24
25
26
27
28
29
1950
1960
1970
1980
1990
2000
2010



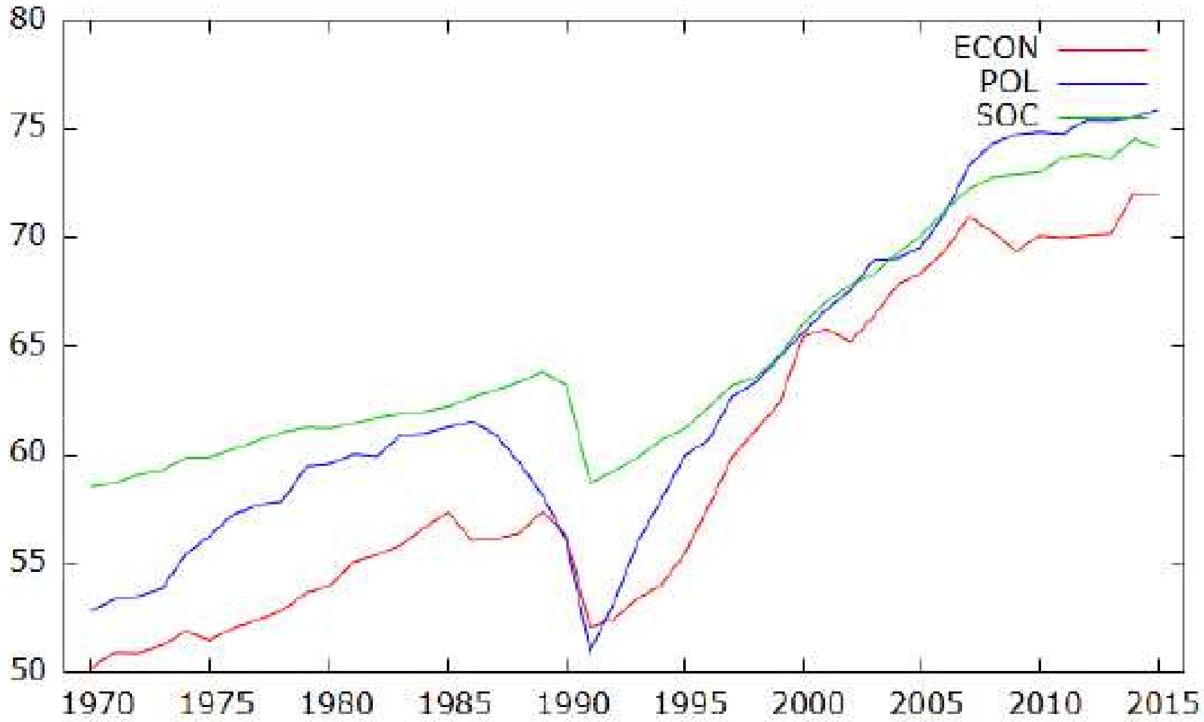
closes
Index
AUDUSD.Close
NZDUSD.Close
0.5
0.6
0.7
0.8
0.9
1.0
1.1
2009
2010
2011
2012
2013
2014
2015

Примеры



Динамика потребления электроэнергии (мегаватт на душу населения) и ВВП (в долларах) на единицу потребления электроэнергии (кг) в России

Примеры



ECON – индекс экономической глобализации;
SOC – индекс социальной глобализации;
POL – индекс политической глобализации.

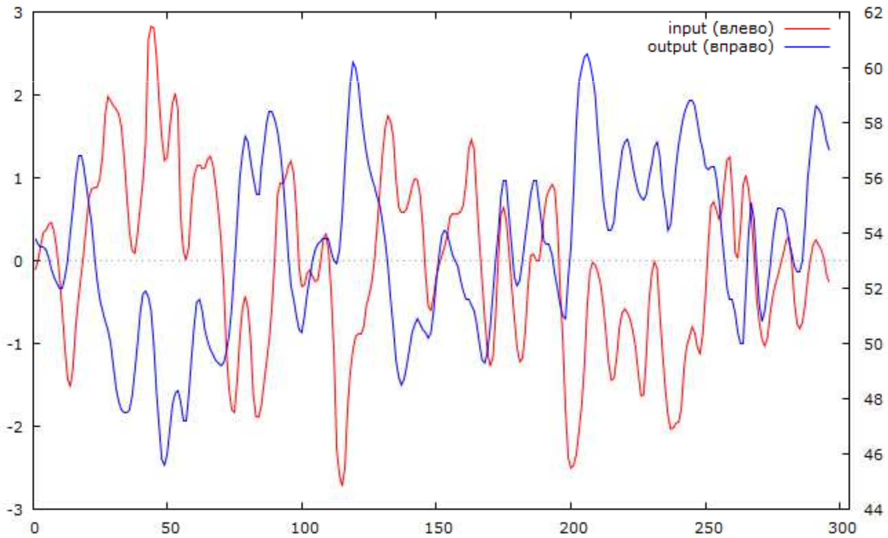
Динамика индексов глобализации

Регрессионный анализ со стационарными ВР

процессы x_t - стационарный, y_t - стационарный

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \quad t = \overline{1,T}, \quad \varepsilon_t \sim WN$$

Пример: (Box, Jenkins, and Reinsel 2008, Series J) on the input and output of a gas furnace, where 296 paired observations on the input (gas rate) and output (% CO2) were recorded every 9 seconds.



-МНК: Применимы классические тесты

- необходимо анализировать остатки

- при наличии детерминированного тренда в одном из ряду:

$$y_t = \beta_1 + \beta_2 x_t + \gamma t + \varepsilon_t, \quad t = \overline{1,T}, \quad \varepsilon_t \sim WN$$

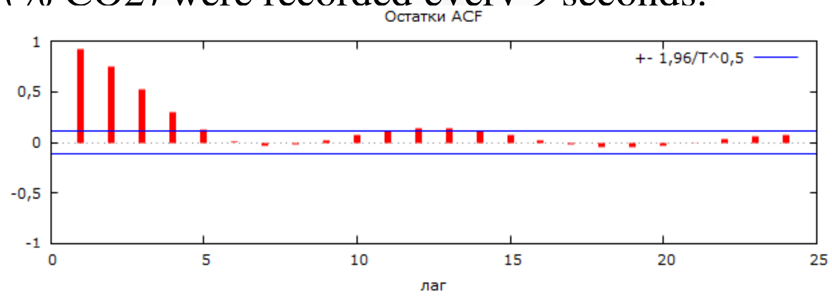
процессы x_t, y_t – стационарны: $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \quad t = \overline{1,T}, \quad \varepsilon_t \sim WN$

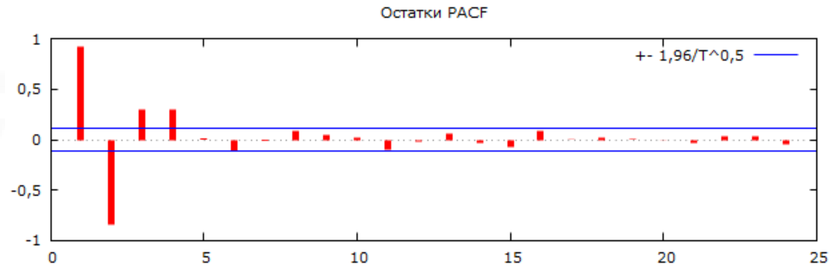
Пример: (Box, Jenkins, and Reinsel 2008, Series J) on the input and output of a gas furnace, where 296 paired observations on the input (gas rate) and output (% CO2) were recorded every 9 seconds.

Зависимая переменная: output

	Коэффициент	Ст. ошибка	z	P-значение	
const	53,4269	0,163317	327,1	0,0000	***
input	-1,44643	0,152292	-9,498	2,15e-021	***

Среднее зав. перемен	53,50912	Ст. откл. зав. перемен	3,202121
Среднее инноваций	0,000000	Ст. откл. инноваций	2,805861
Лог. правдоподобие	-724,3887	Крит. Акаике	1452,777
Крит. Шварца	1460,158	Крит. Хеннана-Куинна	1455,733





Замечание: При наличии в остатках автокорреляции, можно для остатков использовать модель ARMA.

Регрессионный анализ со стационарными ВР

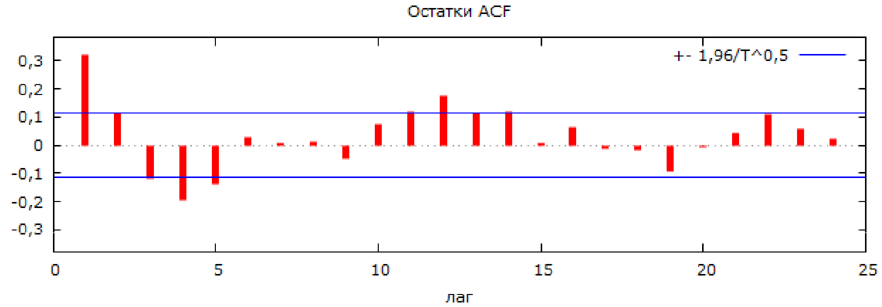
процессы x_t, y_t – стационарны: $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \quad t = \overline{1,T}, \quad \varepsilon_t \sim WN$

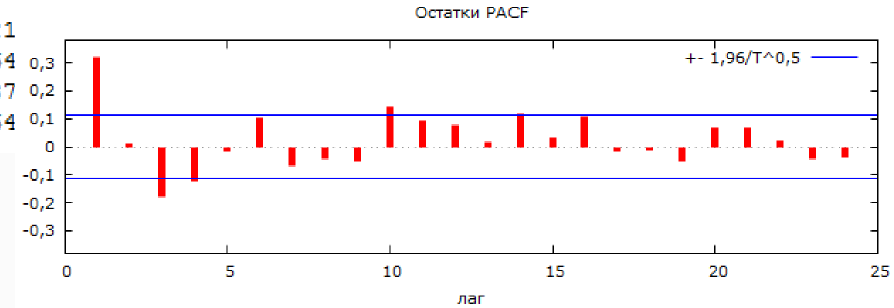
Пример: (Box, Jenkins, and Reinsel 2008, Series J) on the input and output of a gas furnace, where 296 paired observations on the input (gas rate) and output (% CO2) were recorded every 9 seconds.

```
Зависимая переменная: output
Стандартные ошибки рассчитаны на основе Гессиана

             Коэффициент   Ст. ошибка      z       P-значение
  ---------------------------------------------------------
  const        53,5708     0,543378      98,59    0,0000      ***
  phi_1         1,80548    0,0301884     59,81    0,0000      ***
  phi_2        -0,845160   0,0303135    -27,88    4,58e-171   ***
  input         0,445382   0,111059      4,010    6,06e-05    ***

Среднее зав. перемен    53,50912   Ст. откл. зав. перемен   3,202121
Среднее инноваций      -0,000042   Ст. откл. инноваций      0,374364
Лог. правдоподобие     -132,0094   Крит. Акаике             274,0187
Крит. Шварца            292,4705   Крит. Хеннана-Куинна     281,4064
```





Как взаимосвязаны x_t и y_t?

Кросс-корреляционная функция

Кросс-ковариация порядка k для x_t, y_t:

$$\operatorname{cov}\big(y(t), x(t+k)\big) = R_{xy}(k),$$

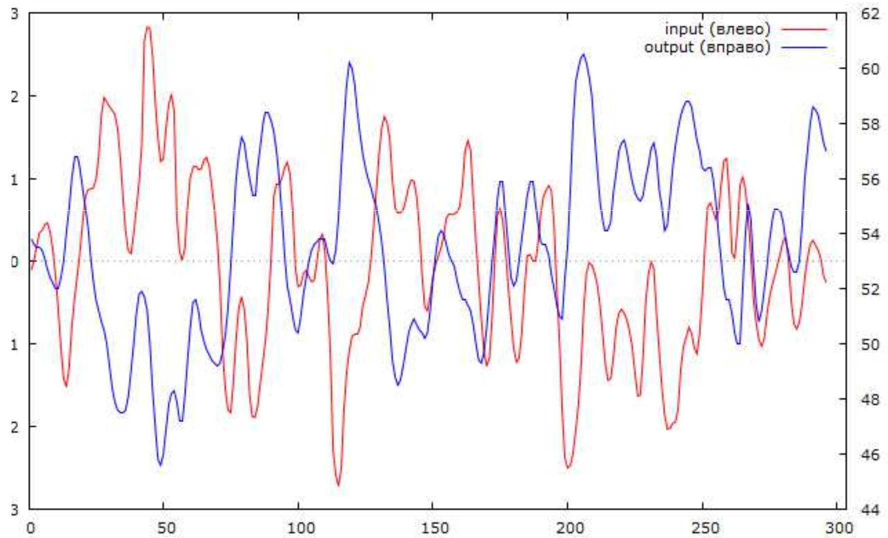
$$R_{xy}(k) \neq R_{xy}(-k),$$

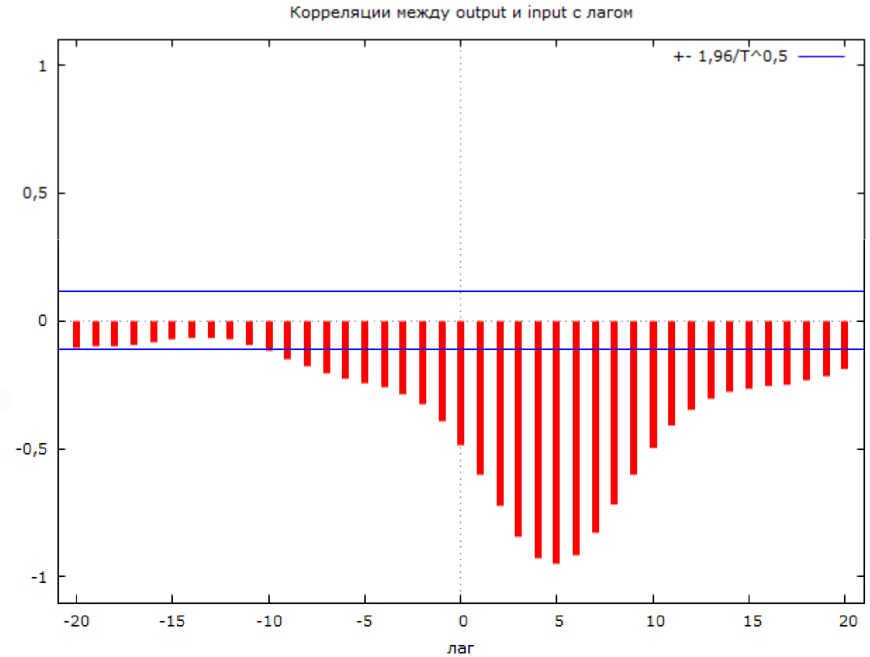
Кросс-корреляция порядка k для x_t, y_t:

$$\rho_{xy}(k) = corr\big(y(t), x(t+k)\big) = \frac{R_{xy}(k)}{\sqrt{R_{xx}(0)R_{yy}(0)}}.$$

Кросс-корреляционная функция

Пример: (Box, Jenkins, and Reinsel 2008, Series J) on the input and output of a gas furnace, where 296 paired observations on the input (gas rate) and output (% CO_2) were recorded every 9 seconds.





http://www.stata.com/manuals13/tsxcorr.pdf

Авторегрессионная модель распределенных лагов

(Autoregressive Distributed Lag Model) ADL (p, q)

$$y_t = \theta + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_q x_{t-q} + \varepsilon_t$$

$$t = \overline{1,T}, \quad \varepsilon_t \sim WN$$

$$\alpha_p(L) y_t = \theta + \beta_q(L) x_t + \varepsilon_t$$

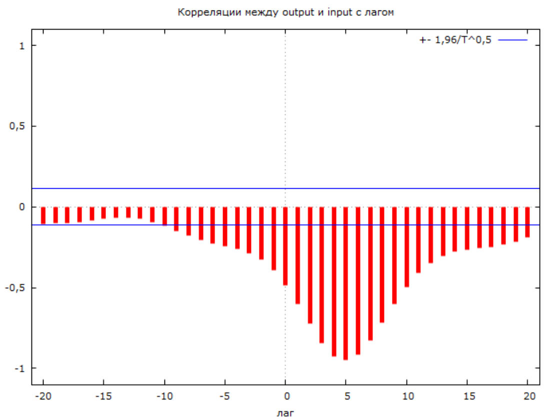
p, q – количество лагов зависимой и независимой переменной

Обобщение ADL: **ARMAX(p, q)**

Случаи: процесс x_t - стационарный, y_t - стационарный

x_t - нестационарный, y_t – нестационарный (р/м позже)

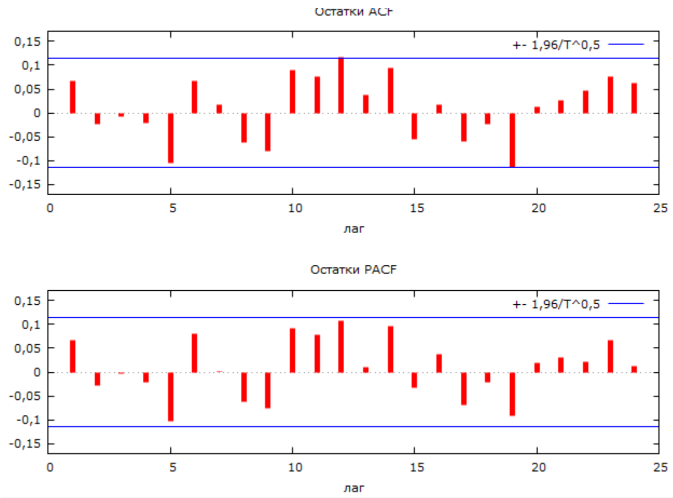
ADL: пример

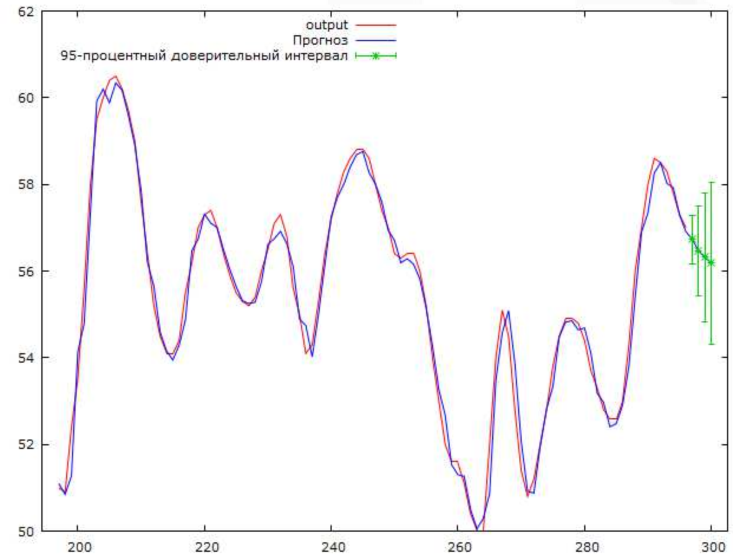



```
Зависимая переменная: output
Стандартные ошибки рассчитаны на основе Гессиана

             Коэффициент   Ст. ошибка      z       P-значение
  ------------------------------------------------------------
  const        53,4779     0,348699      153,4     0,0000      ***
  phi_1         1,60314    0,0523615      30,62    7,30e-206   ***
  phi_2        -0,649716   0,0520851     -12,47    1,03e-035   ***
  input_4      -0,735762   0,0876614      -8,393   4,73e-017   ***
  input_5      -1,02625    0,0923457     -11,11    1,08e-028   ***

Среднее зав. перемен     53,50825   Ст. откл. зав. перемен   3,229551
Среднее инноваций       -0,000898   Ст. откл. инноваций      0,282535
Лог. правдоподобие      -47,09314   Крит. Акаике             106,1863
Крит. Шварца             128,2262   Крит. Хеннана-Куинна     115,0156
```





ADL (1, 1): интерпретация коэффициентов

$$y_t = \theta + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t, t = \overline{1,T}, \quad \varepsilon_t \sim WN$$

$$\alpha_p(L) y_t = \mu + \beta_q(L) x_t + \varepsilon_t$$

- **φ_i - мгновенная реакция, импульсный индикатор (short-run reaction)**: увеличение x_t на единицу влечет непосредственное изменение y_t.

- **Долгосрочный эффект (long-run reaction) (долгосрочный мультипликатор),** при условии стационарности x_t и y_t.

ADL (p, q): интерпретация коэффициентов

$$y_t = \mu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_q x_{t-q} + \varepsilon_t$$

$$t = \overline{1,T}, \quad \varepsilon_t \sim WN$$

$$\alpha_p(L) y_t = \mu + \beta_q(L) x_t + \varepsilon_t \rightarrow$$

$$y_t = (\alpha_p(L))^{-1} \mu + (\alpha_p(L))^{-1} \beta_q(L) x_t + (\alpha_p(L))^{-1} \varepsilon_t,$$

$$y_t = \tilde{\mu} + \sum_{i=0}^{\infty} \varphi_i L^i x_t + \tilde{\varepsilon}_t,$$

- **φ_i - мгновенная реакция, импульсный индикатор (short-run reaction)**: увеличение x_t на единицу влечет непосредственное изменение y_t.

- **Долгосрочный эффект (long-run reaction) (долгосрочный мультипликатор),** при условии стационарности x_t и y_t.

$$\frac{\sum_{i=1}^{q} \beta_i}{1 - \sum_{i=1}^{p} \alpha_i}$$

ADL (p, q): долгосрочный эффект, пример

$$\alpha_p(L)\mathrm{y_t} = \mu + \beta_\mathrm{q}(\mathrm{L})\,\mathrm{x_t} + \varepsilon_t \rightarrow \mathrm{y_t} = \tilde{\mu} + \sum_{\mathrm{i=0}}^{\infty} \varphi_\mathrm{i} \mathrm{L^i}\,\mathrm{x_t} + \tilde{\varepsilon}_t,$$

Рассчитать долгосрочный эффект влияния $\mathrm{x_t}$ на $\mathrm{y_t}$

$$\frac{\sum_{\mathrm{i=1}}^{\mathrm{q}} \beta_\mathrm{i}}{1 - \sum_{\mathrm{i=1}}^{p} \alpha_\mathrm{i}}$$

1. $\mathrm{y_t} = 3 + 1{,}5\,\mathrm{x_t} + 0{,}9\,\mathrm{x_{t-1}} + 0{,}3\,\mathrm{x_{t-2}} + \varepsilon_\mathrm{t}$
2. $\mathrm{y_t} = 0{,}1 + 0{,}2\,\mathrm{x_t} - 0{,}02\,\mathrm{x_{t-1}} + 0{,}6\,\mathrm{y_{t-1}} + \varepsilon_\mathrm{t}$

Представление ADL (1, 1) через ECM

Пусть y_t и x_t - стационарны: $y_t = \theta + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$

y_t – объемы продаж компании, x_t - затраты на рекламу в месяце t.

Преобразование модели (анализ долгосрочного и краткосрочного поведения динамического соотношения)

Замена: $y_t = y_{t-1} + \Delta y_t, x_t = x_{t-1} + \Delta x_t$

$$\Delta y_t = \theta + \beta_0 \Delta x_t - (1-\alpha_1) y_{t-1} + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t$$

$$\boxed{\Delta y_t = \beta_0 \Delta x_t - (1-\alpha_1)\left(y_{t-1} - \frac{\theta}{1-\alpha_1} - \frac{\beta_0+\beta_1}{1-\alpha_1} x_{t-1}\right) + \varepsilon_t}$$

Мгновенный отклик Δy_t на текущее изменение Δx_t

Отклонение от долгосрочного равновесия в момент времени (t-1)

Это представление ADL модели наз. **моделью коррекции ошибками** (error correction model, ECM)

Замечание: $|\alpha_1|<1 \rightarrow (1-\alpha_1)>0 \rightarrow$ коэффициент при II-й части всегда отрицательный. $(1-\alpha_1)$ – скорость коррекции при отклонении от равновесного состояния, $1/(1-\alpha_1)$ – период возврата к равновесному состоянию

Представление ADL (1, 1) через ECM: пример

$$y_t = y_{t-1} + \Delta y_t, x_t = x_{t-1} + \Delta x_t$$

$$\Delta y_t = \theta + \beta_0 \Delta x_t - (1-\alpha_1) y_{t-1} + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t$$

$$\boxed{\Delta y_t = \beta_0 \Delta x_t - (1-\alpha_1)\left(y_{t-1} - \frac{\theta}{1-\alpha_1} - \frac{\beta_0+\beta_1}{1-\alpha_1} x_{t-1}\right) + \varepsilon_t} \qquad \text{ECM}$$

Мгновенный отклик Δy_t на текущее изменение Δx_t

Отклонение от долгосрочного равновесия в момент времени (t-1)

Пример. Представить ADL в виде ECM, дать интерпретацию.

1. $y_t = 0{,}7 + 0{,}5\, y_{t-1} + 0{,}1\, x_t + 0{,}3\, x_{t-1} + \varepsilon_t$
2. $y_t = 2 - 0{,}3\, y_{t-1} + 0{,}1\, x_t + 0{,}9\, x_{t-1} + \varepsilon_t$ (*самостоятельно*)

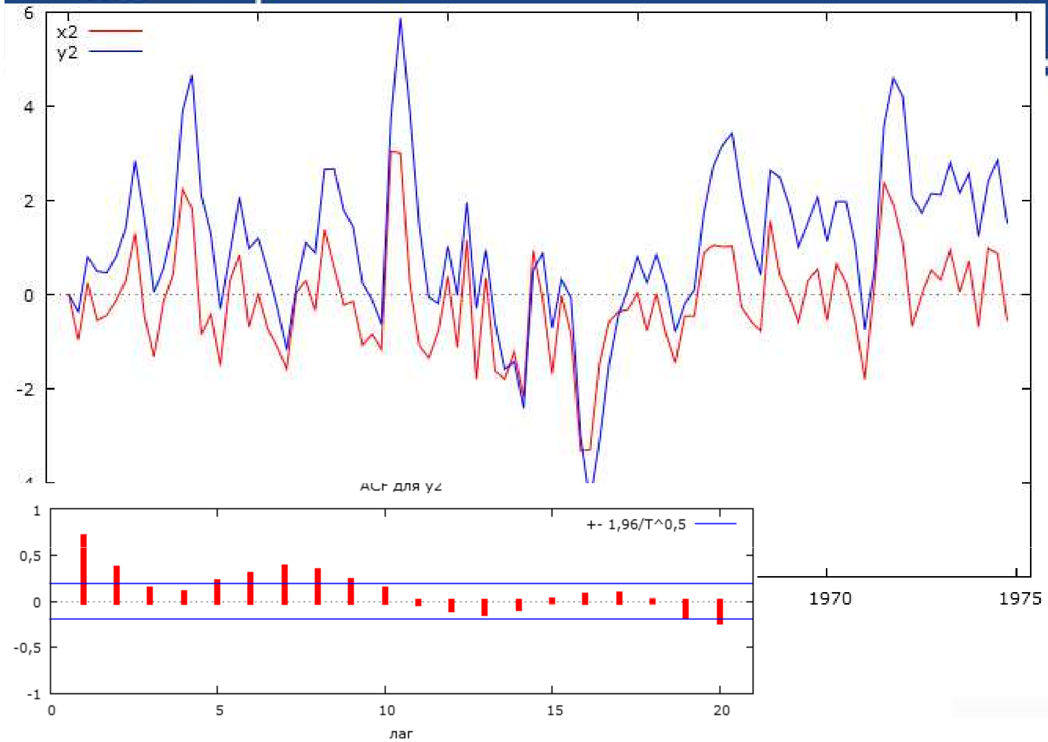
Случай ADL (p, q)

$$\Delta y_t = \beta_0 \Delta x_t + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{i=1}^{q-1} \gamma_i \Delta x_{t-i} - \alpha_p(1)\left(y_{t-1} - \frac{\theta}{\alpha_p(1)} - \frac{\beta_q(1)}{\alpha_p(1)} x_{t-1}\right) + \varepsilon_t$$

- Оценивание и диагностика ADL – аналогично ARIMA
- Метод оценивания МНК → состоятельные оценки
- Все переменные должны быть стационарными
- 2 подхода оценивания
  1. Оценить ADL → пересчитать ECM
  2. Оценить ECM
- Прогнозирование («Квантиль» №1, 2006)
- Примеры

Канторович Г.Г. Лекция №11. Эк журнал ВШЭ (2002)
http://uisrussia.msu.ru/docs/nov/hse_ejournal/2002/4/06_04_06.pdf

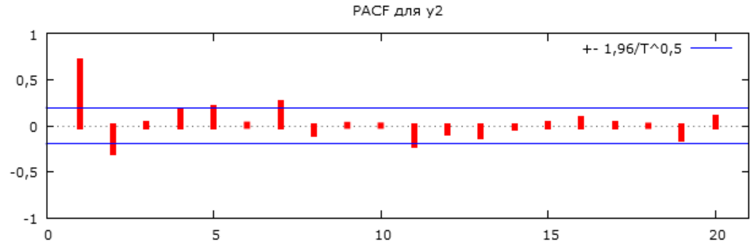
ADL (p, q): пример

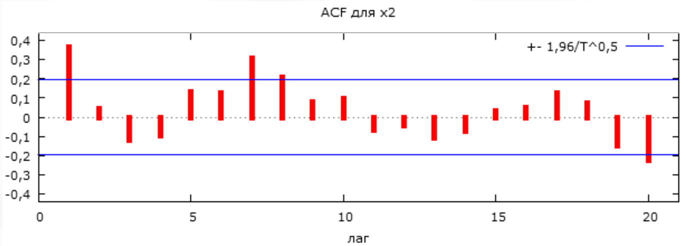


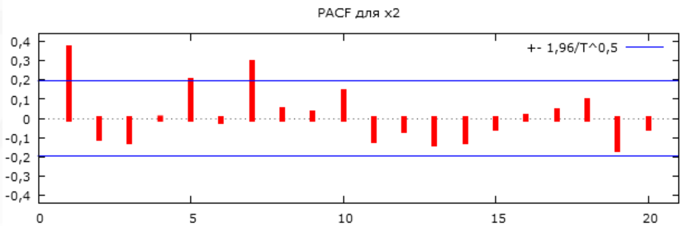
Дано: $x_t = 0.3x_{t-1} + \varepsilon_t$

$$(1-0.5L)y_t = 0.7 + (0.1 + 0.3L)x_t + \varepsilon_t$$

Стационарность процессов:








```
. dfuller  y2
Dickey-Fuller test for unit root                   Number of obs   =        99
```

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
		Interpolated Dickey-Fuller		
Z(t)	**-4.131**	**-3.511**	**-2.891**	**-2.580**

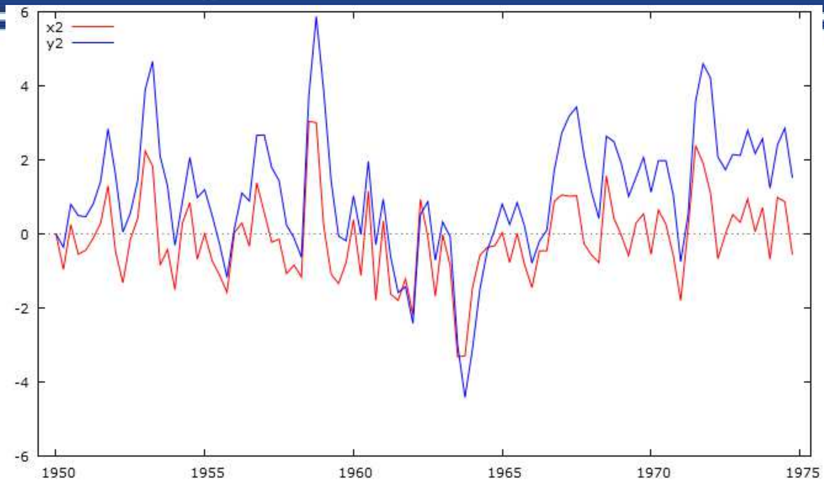
MacKinnon approximate p-value for Z(t) = **0.0009**

```
. dfuller  x
Dickey-Fuller test for unit root                   Number of obs   =        99
```

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
		Interpolated Dickey-Fuller		
Z(t)	**-6.708**	**-3.511**	**-2.891**	**-2.580**

MacKinnon approximate p-value for Z(t) = **0.0000**

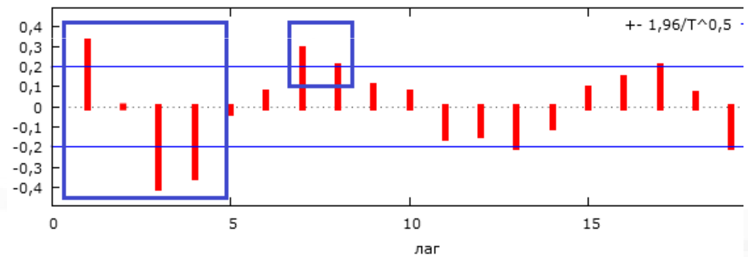
ADL (p, q): пример

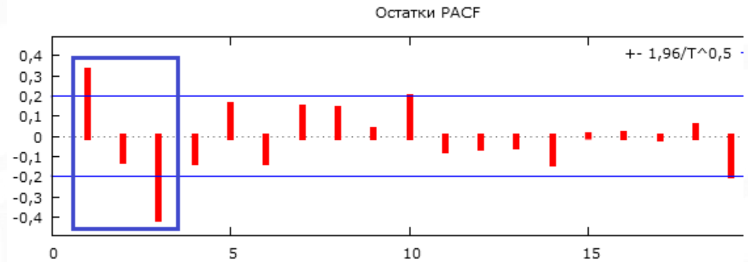


Дано: $x_t = 0.3x_{t-1} + \varepsilon_t$

$$(1-0.5L)y_t = 0.7 + (0.1+0.3L)x_t + \varepsilon_t$$

Подбор модели и анализ остатков.



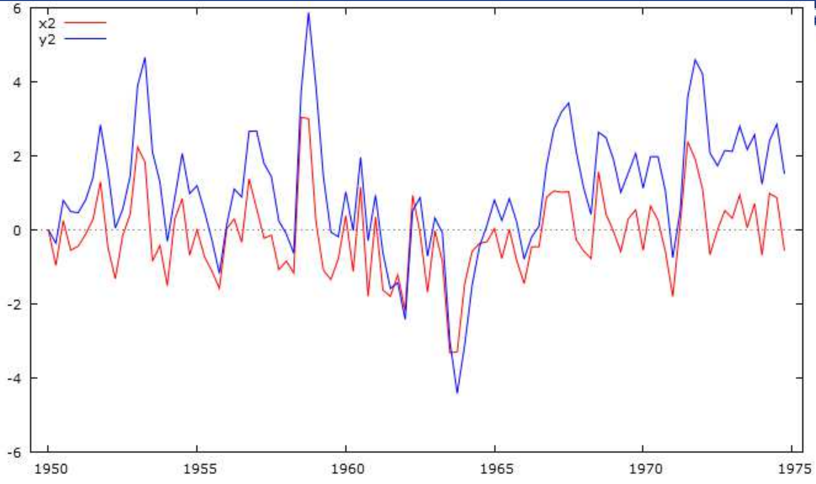


ADL (1, 1)

	Коэффициент	Ст. ошибка	z	Р-значение	
const	1,26723	0,161022	7,870	3,55e-015	***
phi_1	0,839467	0,0557377	15,06	2,92e-051	***
x2	1,01769	0,0239304	42,53	0,0000	***
x2_1	0,437512	0,0243359	17,98	2,89e-072	***

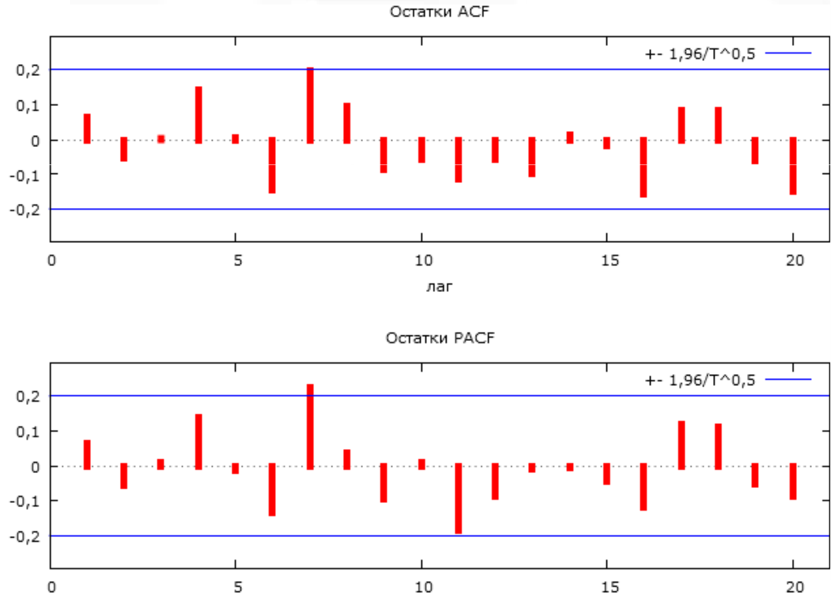
Среднее зав. перемен	1,077766	Ст. откл. зав. перемен	1,695354
Среднее инноваций	0,005479	Ст. откл. инноваций	0,270142
Лог. правдоподобие	-11,51296	Крит. Акаике	33,02593
Крит. Шварца	46,00153	Крит. Хеннана-Куинна	38,27588

ADL (p, q): пример



$$x_t = 0.3x_{t-1} + \varepsilon_t$$
$$(1-0.5L)\,y_t = 0.7 + (0.1+0.3L)\,x_t + \varepsilon_t$$

Подбор модели и анализ остатков.



ADL (6, 2)

	Коэффициент	Ст. ошибка	z	P-значение	
const	1,30978	0,147887	8,857	8,25e-019	***
phi_1	2,36879	0,0862359	27,47	4,15e-166	***
phi_2	-2,59411	0,178994	-14,49	1,35e-047	***
phi_3	1,38190	0,143970	9,599	8,11e-022	***
phi_5	-0,423789	0,0964222	-4,395	1,11e-05	***
phi_6	0,227536	0,0644637	3,530	0,0004	***
x2	1,06418	0,00548716	193,9	0,0000	***
x2_1	0,440550	0,00841800	52,33	0,0000	***
x2_2	0,114650	0,00555207	20,65	9,78e-095	***

Среднее зав. перемен	1,092388	Ст. откл. зав. перемен	1,697784
Среднее инноваций	0,000115	Ст. откл. инноваций	0,066022
Лог. правдоподобие	123,9517	Крит. Акаике	-227,9033
Крит. Шварца	-202,0536	Крит. Хеннана-Куинна	-217,4476

Ljung-Box Q' = 13,0249,
p-значение = P(Хи-квадрат(7) > 13,0249) = 0,07150