Методология Box–Jenkins (BJ) или модели AR(I)MA.

Модель авторегрессии AR(p).

Модель авторегрессии MA(q).

- Целью эконометрического моделирования часто является так называемое out-of-sample предсказание, т.е. прогноз

  гипотеза, предсказание, прогноз, ретро-прогноз, сценарий

  стат. методы, моделирование, экспертные оценки и интуитивные оценки

- Для решения задачи прогнозирования возможно использовать как одномерные, так и многомерные (структурные) факторные эконометрические модели, за основу в которых берутся постулаты экономической теории

- С другой стороны, за основу модели может быть взят принцип инерционности, как свойства большинства временных рядов экономических показателей

  «let the data speak for themselves»

- Авторегрессионный процесс (AR) — одна из основных компонент моделей Бокса-Дженкинса

- Простейший случай: процесс авторегрессии первого порядка, обозначаемый AR(1), определяется как

$$y_t = \delta + \rho \cdot y_{t-1} + \varepsilon_t,$$

где $M(\varepsilon_t) = 0,\ D(\varepsilon_t) = \sigma^2,\ Cov(\varepsilon_t, \varepsilon_{t-s}) = 0 \quad \forall t, s.$

- В модели AR(1), или в авторегрессионной модели первого порядка, все значения временного ряда y_t определяются линейно по y_{t-1}

- δ, ρ, σ^2 — неизвестные параметры, постоянные (т.е. $const$)

- ε_t процесс «белого шума» или Гауссов «белый шум»: гомоскедастичная случайная компонента с нулевой автокорреляцией (которую т.о. невозможно предсказать на основе ее предыдущих значений)

- Для процесса AR(1)

$$M(y_t) = M(\delta + \rho \cdot y_{t-1} + \varepsilon_t) = \delta + \rho \cdot M(y_{t-1}) = \frac{\delta}{1-\rho} \equiv \mu$$

$[M(y_{t-1}) = \delta + \rho \cdot M(y_{t-2});\ M(y_t) = \delta + \rho \cdot M(y_{t-1}) = \delta + \rho \cdot (\delta + \rho \cdot M(y_{t-2}))]$

при предположении, что $M(y_t)$ не зависит от t. Здесь и далее: $|\rho| < 1$.

- Для упрощения часто рассматриваются центрированные процессы: с $cy_t = y_t - \mu$. Процесс AR(1) может быть переписан как

$$cy_t = \rho \cdot cy_{t-1} + \varepsilon_t$$

- Все результаты, полученные для центрированных процессов, можно расширить на исходные процессы добавлением константы

- Для (*центрированного*) процесса AR(1) дисперсия может быть записана как:

$$D(y_t) = D(\rho \cdot y_{t-1} + \varepsilon_t) = \rho^2\ D(y_{t-1}) + D(\varepsilon_t) = \frac{\sigma^2}{1-\rho^2}$$

при предположении, что $D(y_t)$ не зависит от t

- Динамические свойства процесса AR(1) могут быть обобщены при выводе автоковариации:

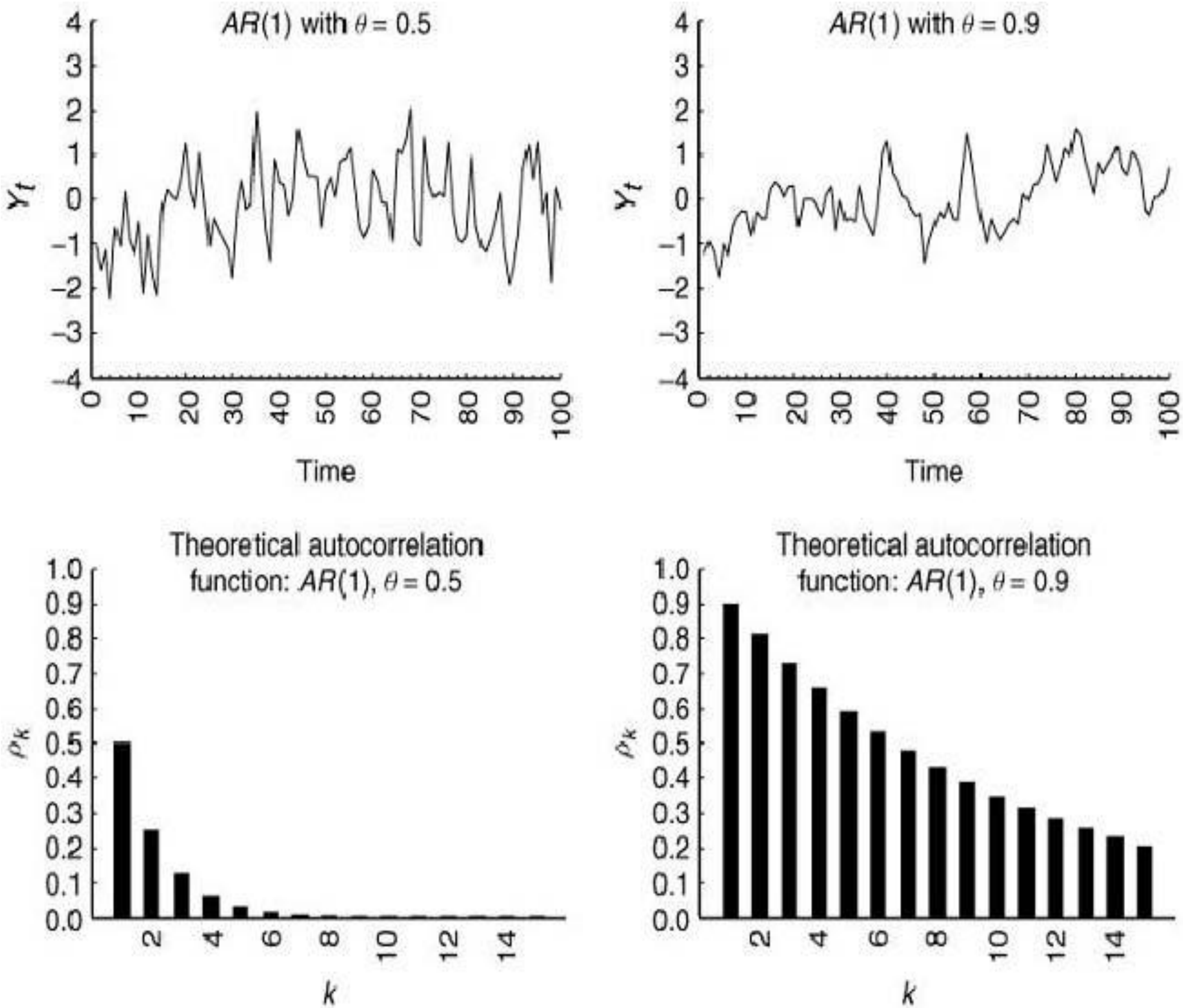
$$Cov(y_t, y_{t-1}) = \frac{1}{n}\sum(y_t - \mu)\cdot(y_{t-1} - \mu) = \frac{1}{n}\sum cy_t \cdot cy_{t-1} = M(cy_t \cdot cy_{t-1}) =$$

$$= M[(\rho \cdot cy_{t-1} + \varepsilon_t)\cdot cy_{t-1}] = \rho \cdot D(y_{t-1}) = \rho\frac{\sigma^2}{1-\rho^2}$$

Таким образом, очевидно, что $Cov(y_t, y_{t-k}) = \rho^k \dfrac{\sigma^2}{1-\rho^2}$ при $k = 1, 2, 3...$

- Свойства корреляции:
  - зависит только от значения k, не зависит от t
  - Убывает с ростом значения k
  - Равна нулю только при $\rho = 0$.

- Допущение $|\rho| < 1$, т.е. предположение о постоянстве предельных моментов, является условием стационарности процесса AR(1).



AR(1) with θ = 0.5
AR(1) with θ = 0.9
Y_t
Time
Theoretical autocorrelation function: AR(1), θ = 0.5
Theoretical autocorrelation function: AR(1), θ = 0.9
ρ_k
k

- Используя понятие центрированного процесса можно записать обобщенную модель AR(p):

$$cy_t = \rho_1 \cdot cy_{t-1} + \rho_2 \cdot cy_{t-2} + \ldots + \rho_p \cdot cy_{t-p} + \varepsilon_t$$

- *Вопрос первый*: критерий стационарности процесса AR(p)?
- *Вопрос второй*: определение порядка p и «цена» использования большого количества параметров (ограниченных условиями стационарности)?

- Введем определение оператора лага *L*: $Lcy_t = cy_{t-1}$
- Оператор L может рассматриваться, как постоянный, т.е. :

$$L^2 cy_t = L(Lcy_t) = Lcy_{t-1} = cy_{t-2}$$

и в общем случае $L^p cy_t = cy_{t-p}$

  - Аналогично для отрицательных значений p: $L^{-1} cy_t = cy_{t+1}$
  - Для константы: $L\mu_t = \mu$

- Используя оператор взятия лага, процесс AR(1) может быть записан как

$$\rho(L) cy_t = \varepsilon_t, \text{ где } \rho(L) = 1 - \rho L$$

- И обобщенный процесс AR(p):

$$\rho(L) cy_t = \varepsilon_t,\ \rho(L) = 1 - \rho_1 L - \rho_2 L^2 - \dots \rho_p L^p$$

- Для процесса AR(1), т.е. при p=1 условие стационарности процесса: $|\rho| < 1$

- Рассмотрим процесс AR(2). Имеем $\rho(L) = 0$

$$1 - \rho_1 L - \rho_2 L^2 = (1 - \theta_1 L)(1 - \theta_2 L),$$ откуда $|\theta_1| < 1$, $|\theta_2| < 1$

- Аналогичный вывод можно получить, используя понятие характеристического уравнения:

$$(1 - \theta_1 z)(1 - \theta_2 z) = 0,$$

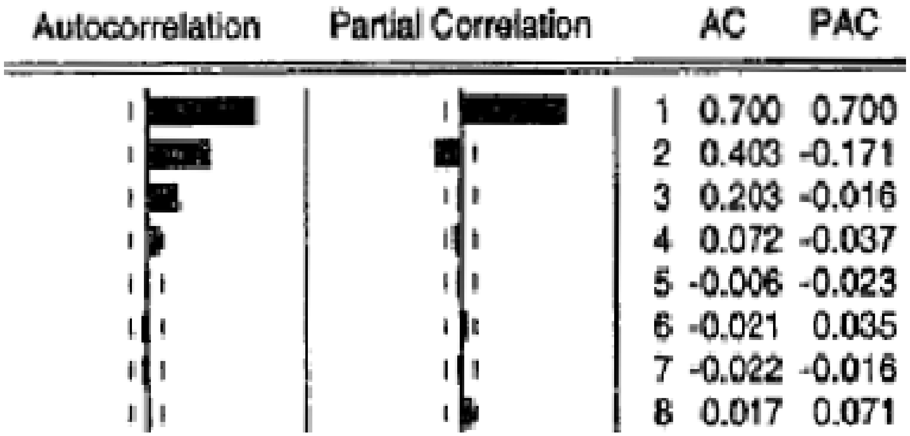
где z_1, z_2 — корни характеристического уравнения. Тогда условие стационарности определяется, как $|\theta_i| < 1;\ |z_i| > 1.$

- Пример AR(2): $cy_t = 1{,}2 \cdot cy_{t-1} - 0{,}32 \cdot cy_{t-2} + \varepsilon_t$

$$1 - \rho_1 L - \rho_2 L^2 = 1 - 1{,}2L + 0{,}32L^2 = (1 - 0{,}8L)(1 - 0{,}4L) = 0,$$

$$(1 - 0{,}8z)(1 - 0{,}4z) = 0 \text{ и } z_1 = 1/0{,}8;\ z_2 = 1/0{,}4.$$

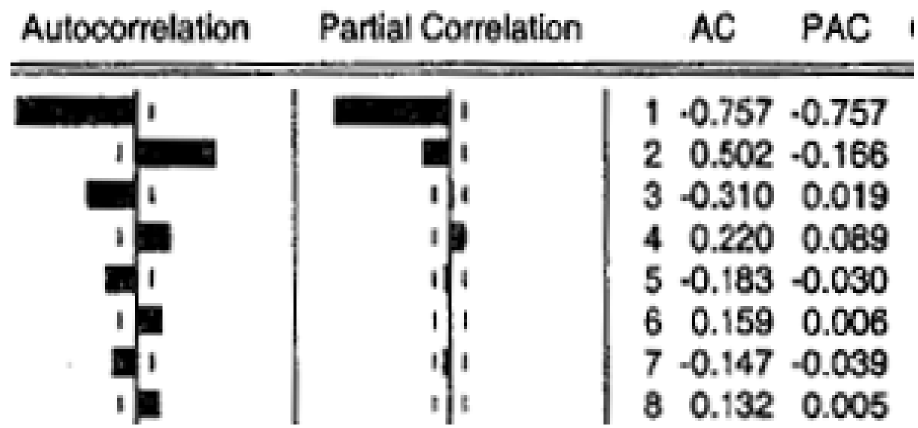
Оба корня характеристического уравнения удовлетворяют условию $|z_i| > 1$.



	AC	PAC
1	0.700	0.700
2	0.403	-0.171
3	0.203	-0.016
4	0.072	-0.037
5	-0.006	-0.023
6	-0.021	0.035
7	-0.022	-0.016
8	0.017	0.071

$y_t = 0{,}8y_{t-1} - 0{,}2y_{t-2} + e_t$, корни

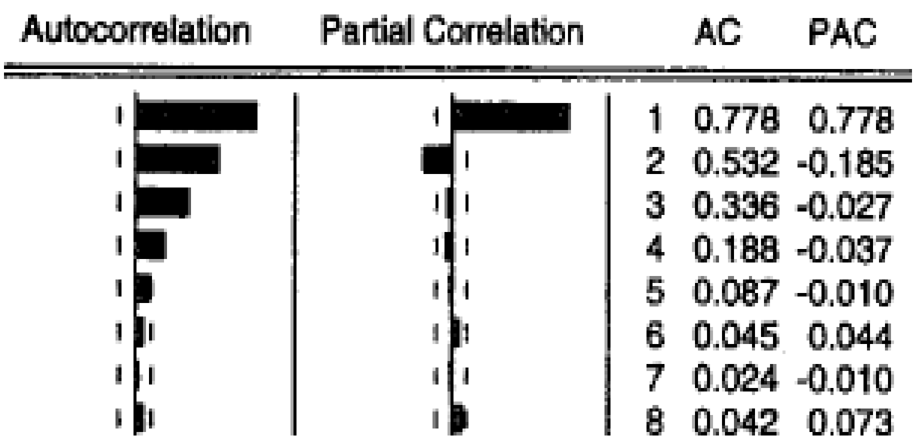
$\rho_{1,2} = 2 \pm i$



	AC	PAC
1	-0.757	-0.757
2	0.502	-0.166
3	-0.310	0.019
4	0.220	0.089
5	-0.183	-0.030
6	0.159	0.006
7	-0.147	-0.039
8	0.132	0.005

$y_t = -0{,}9y_{t-1} - 0{,}2y_{t-2} + e_t$, корни

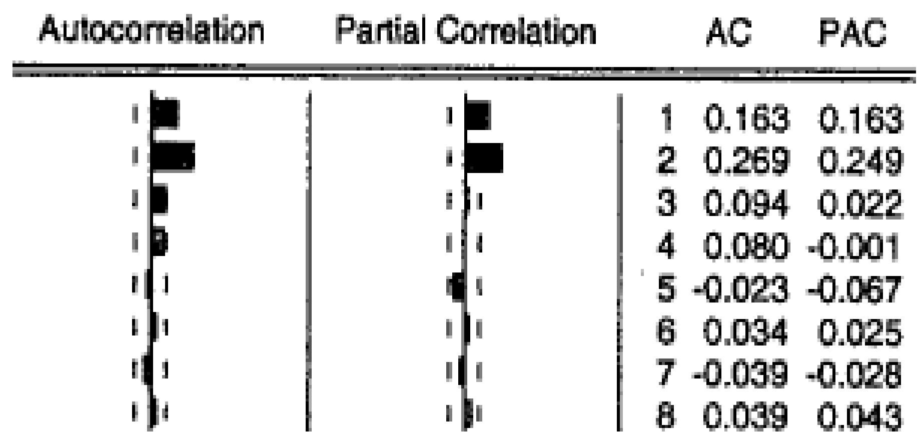
$\rho_{1,2} = \{-2{,}5; -2\}$



	AC	PAC
1	0.778	0.778
2	0.532	-0.185
3	0.336	-0.027
4	0.188	-0.037
5	0.087	-0.010
6	0.045	0.044
7	0.024	-0.010
8	0.042	0.073

$y_t = 0{,}9y_{t-1} - 0{,}2y_{t-2} + e_t$, корни

$\rho_{1,2} = \{2{,}5; 2\}$



	AC	PAC
1	0.163	0.163
2	0.269	0.249
3	0.094	0.022
4	0.080	-0.001
5	-0.023	-0.067
6	0.034	0.025
7	-0.039	-0.028
8	0.039	0.043

$y_t = 0{,}1y_{t-1} + 0{,}2y_{t-2} + e_t$, корни

$\rho_{1,2} = \{-2{,}5; 2\}$

Согласно методологии Бокса-Дженкинса построение модели $AR(p)$ включает в себя следующие этапы:

- Идентификация модели (определение значения p);
- Оценивание параметров модели;
- Анализ адекватности модели;
- Прогнозирование по модели.

Модели Бокса-Дженкинса являются моделями ***стационарных*** временных рядов. Если временной ряд является нестационарным и относится к классу DS-рядов, то прежде его требуется преобразовать, взяв разности так, чтобы он стал стационарным.

- Основное предназначение моделей Бокса-Дженкинса – прогнозирование, что следует учитывать при визуальном анализе временного ряда, в случае, когда ряд явно содержит линейный детерминированный тренд, изломы тренда, сезонные колебания и т.п.

- Если временной ряд содержит в себе детерминированный тренд, то можно говорить о процедуре выделения тренда. В более *примитивном* варианте это означает, что в модель может быть введена переменная тренда.

- Аналогично справедливо для временных рядов с выраженной сезонностью. Модификации моделей Бокса-Дженкинса с учетом сезонностью строятся на основе временных рядов с устраненной сезонной компонентой с помощью оператора взятия сезонной последовательной разности. Например

$$d_4 y_t = \left(1 - L^4\right) y_t = y_t - y_{t-4} \text{ для квартальных данных}$$

$$d_{12} y_t = \left(1 - L^{12}\right) y_t = y_t - y_{t-12} \text{ для помесячных данных}$$

- Этап идентификации модели
  - Визуальный анализ графика временного ряда;
  - Анализ коррелограммы (функций ACF, $PACF$);
  - Тесты «единичного корня».
- Исходя из практических рекомендаций, большинство временных рядов можно аппроксимировать с помощью следующих моделей:
  - *Один параметр (p)*: ACF - экспоненциально убывает; $PACF$ - имеет резко выделяющееся значение для лага *1*, как правило, нет корреляций на других лагах.
  - *Два параметра авторегрессии (p)*: ACF имеет форму синусоиды (экспоненциально убывает в первых лагах); $PACF$ имеет резко выделяющиеся значения на лагах *1*, *2*, нет корреляций на других лагах.

!! Если временной ряд $\{y_t\}$ не стационарен, строим модель $AR(1)$ для первой разности временного ряда: $\Delta y_t = \gamma_0 + \gamma_1 \Delta y_{t-1} + \varepsilon_t$.

- Этап оценивания параметров модели
  - Метод наименьших квадратов;
  - Метод максимального правдоподобия;
  - Метод моментов.
- Этап тестирования адекватности модели
  - Статистическая значимость параметров модели;
  - Нормальное распределение остатков модели;
  - Некоррелированность остатков модели;
  - Стационарность остатков модели («белый шум»);
  - Экономность модели (наиболее простая модель).
- Критерием выбора адекватной модели могут выступать информационные критерии Акайке (AIC) и Шварца (SIC).

- Вторая компонента моделей Бокса-Дженкинса – модель скользящего среднего первого порядка, обозначаемая MA(1), определяется как

$$y_t = \mu + \varepsilon_t + \alpha \cdot \varepsilon_{t-1},$$

где

$$M(y_t) = \mu,$$

$$D(y_t) = (1+\alpha^2)\sigma^2,$$

$$Cov(y_t, y_{t-k}) = \begin{cases} \alpha\,\sigma^2, \textit{если } k = 1 \\ 0, \textit{если } k > 1 \end{cases} \quad \forall t, k$$

- В общем виде MA(q) записывается, как

$$y_t = \mu + \varepsilon_t + \alpha_1 \cdot \varepsilon_{t-1} + \alpha_2 \cdot \varepsilon_{t-2} + \alpha_3 \cdot \varepsilon_{t-3} + ... + \alpha_q \cdot \varepsilon_{t-q},$$

- Всякий стационарный процесс может быть представлен как процесс скользящего среднего MA(∞) — процесс с некоторыми коэффициентами (сумма их модулей должна быть конечной)

  - Т.е. значения любого стационарного временного ряда можно сколь угодно точно приблизить некоторым MA(q)-процессом конечного порядка

  - Для сокращения количества параметров модели, т.е. понижения порядка q модели MA(q) дополняют авторегрессионной частью. Такие модели принято называть ARMA-моделями (или ARIMA-моделями для нестационарных, но интегрированных, временных рядов; I – порядок интегрирования временного ряда y_t)

- Для процесса AR(1) :

$$
\begin{aligned}
y_t &= \delta + \rho \cdot y_{t-1} + \varepsilon_t = \\
&= \delta + \rho \cdot (\delta + \rho \cdot y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\
&= \delta(1+\rho) + \rho^2 \cdot y_{t-2} + \varepsilon_t + \rho \cdot \varepsilon_{t-1} = \\
&= \delta(1+\rho) + \rho^2 \cdot (\delta + \rho \cdot y_{t-3} + \varepsilon_{t-2}) + \varepsilon_t + \rho \cdot \varepsilon_{t-1} = \\
&= \delta\left(1+\rho+\rho^2\right) + \rho^3 \cdot y_{t-3} + \varepsilon_t + \rho \cdot \varepsilon_{t-1} + \rho^2 \cdot \varepsilon_{t-2} = \ldots.. \\
&= \delta\left(1+\rho+\rho^2+\ldots+\rho^n\right) + \rho^n \cdot y_{t-n} + \varepsilon_t + \rho \cdot \varepsilon_{t-1} + \rho^2 \cdot \varepsilon_{t-2} + \ldots + \rho^n \cdot \varepsilon_{t-n}
\end{aligned}
$$

- Для $|\rho| < 1$ (условие стационарности) и при $n \to \infty$ имеем ***представление в виде MA(∞) процесса AR(1)***:

$$
y_t = \delta + \varepsilon_t + \rho \cdot \varepsilon_{t-1} + \rho^2 \cdot \varepsilon_{t-2} + \ldots + \rho^n \cdot \varepsilon_{t-n} + \ldots = \delta + \sum_{j=0}^{\infty} \rho^j \cdot \varepsilon_{t-j}
$$

- Выбор между представлениями в виде AR и MA зависит от целей анализа

Обратимся вновь к оператору взятия лагов L:

процесс AR(1) : $cy_t = \rho \cdot cy_{t-1} + \varepsilon_t \Rightarrow (1-\rho L)cy_t = \varepsilon_t$

$$\rho(L)cy_t = \varepsilon_t, \text{ где } \rho(L) = 1 - \rho L$$

процесс MA(1) : $y_t = \mu + \varepsilon_t + \alpha \cdot \varepsilon_{t-1} \Rightarrow y_t - \mu = \varepsilon_t + \alpha \cdot \varepsilon_{t-1} \Rightarrow cy_t = (1+\alpha L)\varepsilon_t$

$$cy_t = \alpha(L)\varepsilon_t, \text{ где } \alpha(L) = 1 + \alpha L$$

процесс AR(p) : $\rho(L)cy_t = \varepsilon_t$, где $\rho(L) = 1 - \rho_1 L - \rho_2 L^2 - \ldots \rho_p L^p$

процесс MA(q) : $cy_t = \alpha(L)\varepsilon_t$, где $\alpha(L) = 1 + \alpha_1 L + \alpha_2 L^2 + \ldots + \alpha_q L^q$

- Обратный оператор $\alpha(L)\alpha^{-1}(L)=1$
- Если существует обратный оператор $\alpha^{-1}(L)$, то: $\alpha^{-1}(L)cy_t = \varepsilon_t$

  В общем случае запись определяет процесс AR(1) с бесконечным числом лагов

  - Действительно:

$$\rho(L)=1-\rho L \Rightarrow \rho^{-1}(L)=\frac{1}{1-\rho L}=1+\rho L+\rho^2 L^2+\ldots=\sum_{j=0}^{\infty}\rho^j L^j$$

  Т.о. при конечном числе порядка p, обратный оператор лага бесконечен.

  - С помощью обратного оператора процесс AR может быть записан в виде MA

процесса:

$$(1-\rho L)cy_t=\varepsilon_t \Rightarrow (1-\rho L)^{-1}(1-\rho L)cy_t=(1-\rho L)^{-1}\varepsilon_t=\sum_{j=0}^{\infty}\rho^j L^j\varepsilon_t=\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-j}$$

- При определенных допущениях верно и обратное: MA(1) процесс может быть записан в виде авторегрессии:

$$\alpha(L)=1+\alpha L \Rightarrow \alpha^{-1}(L)=\frac{1}{1+\alpha L}=1+(-\alpha)L+(-\alpha)^2L^2+\ldots=\sum_{j=0}^{\infty}(-\alpha)^j L^j$$

$$(1+\alpha L)^{-1}=\sum_{j=0}^{\infty}(-\alpha)^j L^j \text{ при условии } |\alpha|<1$$

$$cy_t=(1+\alpha L)\varepsilon_t \Rightarrow (1+\alpha L)^{-1}cy_t=(1+\alpha L)^{-1}(1+\alpha L)\varepsilon_t \Rightarrow (1+\alpha L)^{-1}cy_t=\varepsilon_t \Rightarrow$$

$$\sum_{j=0}^{\infty}(-\alpha)^j L^j cy_t=\varepsilon_t \Rightarrow \sum_{j=0}^{\infty}(-\alpha)^j cy_{t-j}=\varepsilon_t$$

$$cy_t=\alpha(L)\varepsilon_t \Rightarrow cy_t=\alpha\varepsilon_{t-1}+\varepsilon_t=\alpha\sum_{j=0}^{\infty}(-\alpha)^j cy_{t-j-1}+\varepsilon_t$$

- В случае процесса MA(q) обратимость требуется для большей точности выводов и прогнозов, в случае процесса AR(p) — для стационарности. Отсюда, проверка обратимости процесса аналогична проверке стационарности.

� В общем случае модель авторегрессии и скользящего среднего **ARMA(p,q)** записывается:

$$cy_t = \rho_1 \cdot cy_{t-1} + \rho_2 \cdot cy_{t-2} + ... + \rho_p \cdot cy_{t-p} + \varepsilon_t + \alpha_1 \cdot \varepsilon_{t-1} + \alpha_2 \cdot \varepsilon_{t-2} + ... + \alpha_q \cdot \varepsilon_{t-q}$$

На языке оператора L: $\rho(L)cy_t = \alpha(L)\varepsilon_t$

� Выбор между одной из форм представления процесса зависит от ограничений относительно количества параметров и других аспектов: целью моделирования является прогнозирование (AR), для экстраполяции и описания свойств (MA) и т.д.

� Ключевым является выбор представления с меньшим количеством лагов, т.е. параметров, в новом представлении.

� Интегрированная модель авторегрессии и скользящего среднего **ARIMA(p,i,q)**:

$$d^i cy_t = \rho_1 \cdot d^i cy_{t-1} + \rho_2 \cdot d^i cy_{t-2} + ... + \rho_p \cdot d^i cy_{t-p} + \varepsilon_t + \alpha_1 \cdot \varepsilon_{t-1} + \alpha_2 \cdot \varepsilon_{t-2} + ... + \alpha_q \cdot \varepsilon_{t-q}$$

� Процесс на примере AR(2) из предыдущей лекции — обратим:

$$cy_t = 1{,}2 \cdot cy_{t-1} - 0{,}32 \cdot cy_{t-2} + \varepsilon_t$$

$$1 - \rho_1 L - \rho_2 L^2 = 1 - 1{,}2L + 0{,}32L^2 = (1 - 0{,}8L)(1 - 0{,}4L) = 0$$

� Пример ARMA(2,1) — необратим и нестационарен:

$$cy_t = 1{,}2 \cdot cy_{t-1} - 0{,}2 \cdot cy_{t-2} + \varepsilon_t - 0{,}5 \cdot \varepsilon_{t-1}$$

$$\rho(L) cy_t = \alpha(L) \varepsilon_t$$

$$cy_t - 1{,}2 \cdot cy_{t-1} + 0{,}2 \cdot cy_{t-2} = \varepsilon_t - 0{,}5 \cdot \varepsilon_{t-1}$$

$$(1 - 0{,}2L)(1 - L) cy_t = (1 - 0{,}5L) \varepsilon_t$$

Модель должна иметь спецификацию ARIMA(2,1,1):

$$(1 - 0{,}2L)(1 - L) cy_t = (1 - 0{,}5L) \varepsilon_t \Rightarrow (1 - 0{,}2L) dcy_t = (1 - 0{,}5L) \varepsilon_t$$

$$dcy_t = 0{,}2 \cdot dcy_{t-1} + \varepsilon_t - 0{,}5 \cdot \varepsilon_{t-1}$$

Т.о., при наличии одного «единичного корня», степень обратимого полинома (оператора L) равняется $p - 1$, а $1 - L \equiv d$, где d — оператор взятия разностей.

- При использовании моделей ARMA, в случае, если у процессов AR и MA есть общие корни — их возможно сократить. В таком случае процесс ARMA(p,q) трансформируется в процесс ARMA(p-1,q-1).

- Пример ARMA(2,1) — обратим, стационарен и преобразуется в AR(1) (ARMA(1,0)):

$$cy_t = 1 \cdot cy_{t-1} - 0{,}25 \cdot cy_{t-2} + \varepsilon_t - 0{,}5 \cdot \varepsilon_{t-1}$$

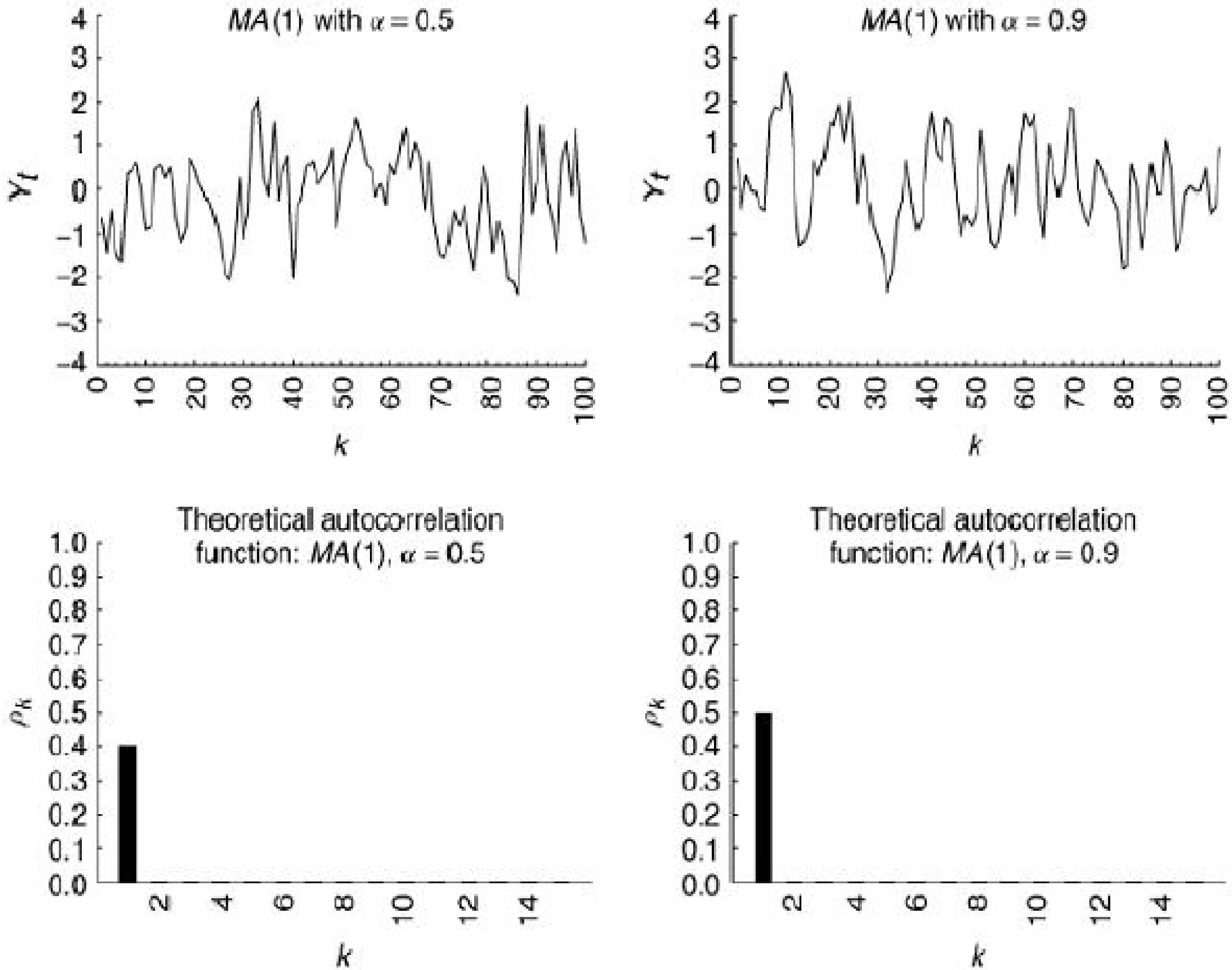
$$\rho(L) cy_t = \alpha(L) \varepsilon_t$$

$$cy_t - 1 \cdot cy_{t-1} + 0{,}2 \cdot cy_{t-2} = \varepsilon_t - 0{,}5 \cdot \varepsilon_{t-1}$$

$$(1 - 0{,}5L)(1 - 0{,}5L) cy_t = (1 - 0{,}5L) \varepsilon_t$$

$$(1 - 0{,}5L)\, cy_t = \varepsilon_t$$

- На практике, для процессов ARMA высоких порядков корни не могут быть тождественно равны, но могут быть достаточно близки. В этом случае модель более низкого порядка имеет меньше параметров и экстраполирует данные почти так же хорошо, т.е. не хуже, как и модель более высокого порядка.


MA(1) with α = 0.5
MA(1) with α = 0.9
Y_t
k
Theoretical autocorrelation function: MA(1), α = 0.5
Theoretical autocorrelation function: MA(1), α = 0.9
ρ_k