Анализ временных рядов и прогнозирование

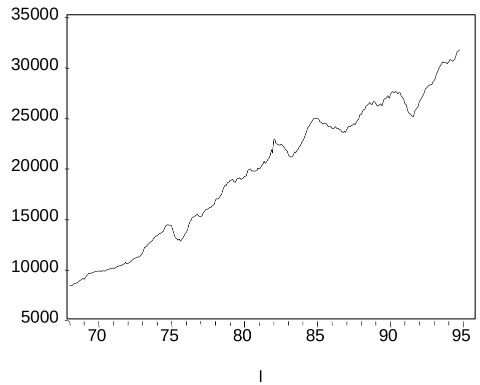
Динамические регрессионные модели

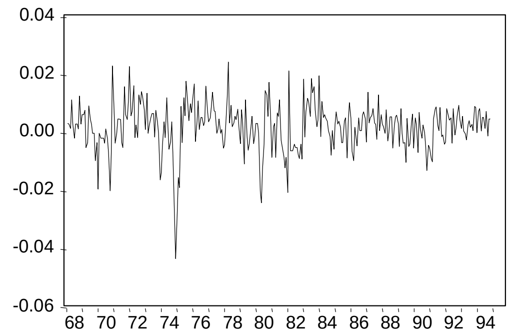
Анализ временных рядов и прогнозирование

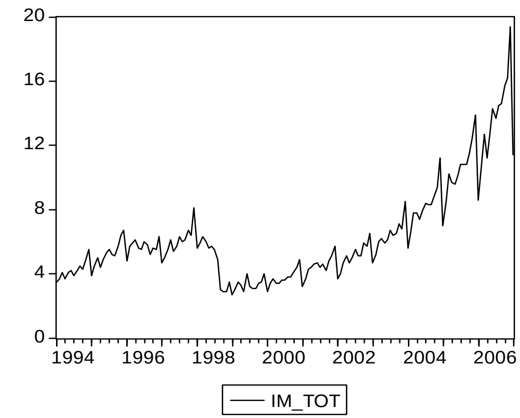
Определение

- Временной ряд (ВР) – это последовательность значений, описывающих протекающий во времени процесс, измеренных в последовательные моменты времени, обычно через равные промежутки.
- Данные типа временных рядов широко распространены в самых разных областях человеческой деятельности. В экономике это ежедневные цены на акции, курсы валют, еженедельные и месячные объемы продаж, годовые объемы производства и т.п.

Графики временных рядов







- на 1-ом графике виден явный линейный тренд,
- на 2-ом – случайные колебания с кластерами волатильности
- на 3-ем – сложный цикл

В рамках нашего курса мы рассмотрим:

- методы подбора математической модели для описания ВР;
- методы выявления периодической и других составляющих ВР;
- прогнозирование поведения ВР ;
- способы изучения взаимозависимостей ВР

T_t

При анализе ВР принято выделять 4 компоненты

- тренд (T)
- циклическая компонента (C)
- сезонная компонента (S)
- случайная компонента (e) – остается после полного вычленения закономерных компонент.

Тренд

- тренд (T) – плавно изменяющаяся компонента, описывающая чистое влияние долговременных факторов (рост население, изменение структуры возрастного состава и т.д.)

Циклическая компонента

- циклическая компонента (С) – плавно изменяющаяся компонента, описывающая длительные периоды относительного подъема и спада,
- состоит из циклов, меняющихся по амплитуде и протяженности
- в экономике бывает связана со взаимодействием спроса и предложения, ростом и истощением ресурсов, изменением в финансовой и налоговой политике и т.п.

Сезонная компонента

- сезонная компонента (S) – состоит из последовательности почти повторяющихся циклов
- объем продаж накануне Нового Года, объем перевозок пассажиров городским транспортом

Случайная компонента

- случайная компонента (е) – остается после полного вычленения закономерных компонент
- иногда содержит часть, поддающуюся моделированию и прогнозированию

ВР представляется

- либо суммой этих компонент

  $$Y=T+C+S+e$$

  в аддитивной модели,
- либо произведением

  $$Y=T*C*S*e$$

  в мультипликативной модели.

Второй вариант более распространен в экономических приложениях и сводится к первому логарифмированием

Трендовая составляющая выделяется методом наименьших квадратов

Наиболее распространены следующие модели трендов:

- линейная - $T_t = \beta_0 + \beta_1 t$
- полиномиальная - $T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_k t^k$
- экспоненциальная - $T_t = \exp(\beta_0 + \beta_1 t)$

Представление о характере тренда можно получить из графика ВР

Учет сезонности

- В аддитивной модели учесть сезонность можно с помощью фиктивных переменных
- Например, уравнение для учета тренда и квартальной сезонности может быть таким:

$$Y_t = \beta_0 + \beta_1 t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \varepsilon_t$$

Построение прогноза для ВР

- Прогнозирование в бизнесе играет очень большую роль, поскольку является рациональной основой для принятия решений.
- Например, предсказание ежемесячных объемов продаж товара – это основа политики контролирования запасов, предсказание будущих доходов корпорации – основа для принятия решений в инвестиционной политике.
- Задача прогнозирования состоит в том, чтобы по имеющимся наблюдениям ВР предсказать его неизвестные будущие значения

Подход Бокса-Дженкинса ARIMA

Методы ARIMA предназначены для моделирования и прогнозирования

- либо стационарных ВР,
- либо ВР, которые могут быть преобразованы к стационарным

Понятие стационарности ВР

- Например, если ВР Y содержит линейный тренд и квартальную сезонность,
- то остатки регрессионной модели

$$Y_t = \beta_0 + \beta_1 t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \varepsilon_t$$

будут стационарным ВР

Понятие стационарности ВР

- Ряд является ***стационарным, если он совершает колебания вокруг своего математического ожидания, которое является константой.***
- Амплитуда колебаний при этом примерно постоянна.
- Сами значения ряда не являются, как правило, независимыми, ***но корреляция между членами ряда*** ε_t ***и*** ε_{t-s} ***зависит только от расстояния*** **s** ***между ними.***

Понятие строгой стационарности

- Ряд ε_t называется строго стационарным или стационарным в узком смысле, если совместное распределение

$$F\left(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T\right)$$

не зависит от сдвига по времени,

т.е. совпадает с распределением

$$F\left(\varepsilon_{1+s}, \varepsilon_{2+s}, \ldots, \varepsilon_{T+s}\right)$$

для любых T и s

Понятие слабой стационарности

- Ряд ε_t называется слабо стационарным
  или стационарным в широком смысле, если

$$E(\varepsilon_t) = \mu = const$$

$$V(\varepsilon_t) = \sigma^2 = const$$

$$\operatorname{cov}(\varepsilon_t, \varepsilon_{t-s}) = \rho(s)$$

- Если нарушается хотя бы одно из этих условий, то ряд называется нестационарным
- Строгая стационарность подразумевает слабую стационарность

TS ряды

- Временной ряд Y_t называется стационарным относительно детерминированного тренда f (t) , если ряд $Y_t - f(t)$ стационарный.
- Тогда говорят, что этот ряд является TS рядом (TS –trend stationary).

Подход Бокса-Дженкинса ARIMA

- Предположение о том, что существует связь между соседними значениями ВР и составляет основу методов ARIMA
- Именно эта гипотеза позволяет предсказать значения $\varepsilon_{t+1}, \varepsilon_{t+2}, \ldots$ на основании известных значений $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_t$
- затем по регрессионной модели

$$Y_t = \beta_0 + \beta_1 t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \varepsilon_t$$

- строятся будущие значения $Y_{t+1}, Y_{t+2}, \ldots$

ARIMA(p,d,q), подбор параметра d

- Ряд ε_t считается реализацией случайного процесса ARIMA(p,d ,q), и в зависимости от типа этого процесса строится прогноз.
- Параметр d называется порядком интегрированности ряда
- Если исходный ряд стационарен, то d=0
- Тогда $\varepsilon_t \sim ARIMA(p,0,q) = ARMA(p,q)$
- На следующем шаге подбираются параметры p и q для процессов AR и MA.

ARIMA(p,d,q), подбор параметра d

- Если исходный ряд не является стационарным, то иногда его можно преобразовать к стационарному виду
- Тогда возможны 2 случая:
  - ряд стационарен в первых разностях

$$\Delta\varepsilon_t = \varepsilon_t - \varepsilon_{t-1} \sim ARMA(p,q)$$

тогда d=1 и $\varepsilon_t \sim ARIMA(p,1,q)$

  - ряд стационарен во вторых разностях

$$\Delta^2\varepsilon_t = \Delta\varepsilon_t - \Delta\varepsilon_{t-1} = \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \sim ARMA(p,q)$$

тогда d= 2 и $\varepsilon_t \sim ARIMA(p,2,q)$

ARIMA(p,d,q), подбор параметра d

- Ряд $\varepsilon_t \sim ARIMA(p,1,q)$ называется интегрированным 1-го порядка $I(1)$
- Ряд $\varepsilon_t \sim ARIMA(p,2,q)$ называется интегрированным 2-го порядка $I(2)$
- Ряды класса TS не относят к интегрированным
- $\varepsilon_t \sim I(0)$ соответствует стационарному ряду, который не является результатом дифференцирования TS ряда

Прогноз процесса AR(p)

Если ε_t ~AR(p), то

$$\varepsilon_t = a_0 + a_1\varepsilon_{t-1} + \ldots + a_p\varepsilon_{t-p} + u_t$$

$$u_t \sim N\left(0, \sigma^2\right)$$

Тогда прогнозное значение

$$\widehat{\varepsilon}_{t+1} = \widehat{a}_0 + \widehat{a}_1\varepsilon_t + \ldots + \widehat{a}_p\varepsilon_{t-p+1}$$

Прогноз процесса MA(q)

Если ε_t ~MA(q), то

$$\varepsilon_t = a_0 + b_1 u_{t-1} + \ldots + b_q u_{t-q} + u_t$$

$$u_t \sim N\left(0, \sigma^2\right)$$

и прогноз

$$\widehat{\varepsilon}_{t+1} = \widehat{a}_0 + \widehat{b}_1 u_t + \ldots + \widehat{b}_q u_{t-q+1}$$

Прогноз процесса ARMA(1,1)

- Если ε_t ~ARMA(1,1), то

$$\varepsilon_t = a_0 + a_1\varepsilon_{t-1} + b_1 u_{t-1} + u_t$$

$$u_t \sim N\left(0, \sigma^2\right)$$

- и прогнозное значение

$$\widehat{\varepsilon}_{t+1} = \widehat{a}_0 + \widehat{a}_1\varepsilon_t + \widehat{b}_1 u_t$$

Алгоритм подбора модели ВР

1. Тестирование на стационарность (тест Unit Root). Если результат положительный, то пункт 3.
2. Приведение к стационарному виду взятием 1-ой или 2-ой разности и снова пункт 1.
3. Идентификация параметров p и q процесса ARMA(p,q) по коррелограммам AC и PAC.
4. Оценивание параметров методом максимального правдоподобия и выбор наилучшей модели (**критерии Акаике, Шварца**).
5. Диагностическая проверка (анализ коррелограмм AC и PAC).

Тестирование стационарности

Тестирование стационарности ВР (тест Дики-Фуллера)

- Для одной из моделей (какой именно, можно выбрать, используя опции)

$$\varepsilon_t = \rho\,\varepsilon_{t-1} + u_{1t}$$

$$\varepsilon_t = \alpha + \rho\,\varepsilon_{t-1} + u_{2t}$$

$$\varepsilon_t = \alpha + \rho\varepsilon_{t-1} + ct + u_{3t}$$

- Оценивается уравнение $\Delta\varepsilon_t = \gamma\,\varepsilon_{t-1} + u_t, \quad \gamma = \rho - 1$
- Проверяется гипотеза $H_0: \quad \gamma = 0$
- Если она не отвергается, это говорит о наличии единичного корня $\rho = 1$, т.е. о нестационарности ВР
- В этом случае $\varepsilon_t = \varepsilon_{t-1} + u_t$ и $\widehat{\varepsilon}_{t+1} = \varepsilon_t$

Тестирование стационарности ВР (тест Дики-Фуллера с константой и трендом)

```
dfuller je, trend regress lags(0)
Dickey-Fuller test for unit root                    Number of obs   =   139
                     ---------- Interpolated Dickey-Fuller ---------
        Test          1% Critical         5% Critical        10% Critical
     Statistic           Value               Value               Value
--------------------------------------------------------------------------
 Z(t)    -1.040          -4.027              -3.445              -3.145
--------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.9384
--------------------------------------------------------------------------
D.je     |      Coef.   Std. Err.      t     P>|t|     [95% Conf. Interval]
---------+----------------------------------------------------------------
      je |
     L1. |  -.0053434  .0051367   -1.04   0.300   -.0155015   .0048147
  _trend |   .0001407  .000036     3.90   0.000    .0000694   .000212
   _cons |  -.0011287  .0011835   -0.95   0.342   -.003469    .0012117
---------------------------------------------------------------------------
```

- Тестовая статистика z(t)= - 1.04, это значение не попадает в критическую область при любом разумном уровне значимости, гипотеза о наличии единичного корня не отвергается, исследуемый ряд нестационарен

Тестирование стационарности ВР (тест Дики-Фуллера)

- Статистика теста (DF-статистика) — это обычная t-статистика для проверки значимости коэффициентов линейной регрессии. Однако, распределение данной статистики отличается от классического распределения.
- Распределение DF-статистики выражается через винеровский процесс и называется распределением Дики — Фуллера.

Тест Дики-Фуллера

- Для каждой из трёх тестовых регрессий существуют свои критические значения *DF*-статистики, которые берутся из специальной таблицы Дики — Фуллера (МакКиннона).
- Если значение статистики лежит левее критического значения (критические значения — отрицательные) при данном уровне значимости, то нулевая гипотеза о единичном корне отклоняется и процесс признается стационарным (в смысле данного теста). В противном случае гипотеза не отвергается и процесс может содержать единичные корни, то есть быть нестационарным (интегрированным) временным рядом.

Тестирование стационарности ВР (тест Дики-Фуллера)

Критические значения статистики Дики — Фуллера при 1%-ном уровне значимости. Для сравнения: критическое значение распределения Стьюдента для всех моделей на больших объёмах выборки — 2,33, на малых выборках — 2,5. МакКинноном выведены приблизительные формулы для оценки критических значений.

Размер выборки	AR-модель	AR-модель с константой	AR-модель с константой и трендом
25	-2,66	-3,75	-4,38
50	-2,62	-3,58	-4,15
100	-2,60	-3,51	-4,04
∞	-2,58	-3,43	-3,96

Теоретически обоснованным является тестирование в первую очередь вторых разностей ряда. Если гипотеза единичного корня для этого ряда отвергается, то тогда тестируется единичный корень в первых разностях. Если на этом этапе гипотеза не отвергается, то исходный ряд имеет два единичных корня. Если отвергается, то проверяется единичный корень в самом временном ряде.

Проблемы теста DF

- Низкая мощность: часто не отвергается исходная (нулевая) гипотеза, когда она в действительности не выполняется
- Невыполнение теоретических предпосылок для вспомогательных моделей в тестах: смещена статистика теста.
- Может даже отвергаться нулевая гипотеза, когда в действительности она верна
- Рассмотренный DF тест рекомендовано использовать при условии гомоскедастичности и некоррелированности случайных отклонений тестируемой модели

Альтернативные тесты

- Расширенный тест DF- ADF для учета автокорреляции

$$\Delta\varepsilon_t = \gamma\,\varepsilon_{t-1} + \sum_{k=1}^{K}\alpha_k \Delta\varepsilon_{t-k} + u_{1t}$$

$$\Delta\varepsilon_t = \alpha + \gamma\,\varepsilon_{t-1} + \sum_{k=1}^{K}\alpha_k \Delta\varepsilon_{t-k} + u_{2t}$$

$$\Delta\varepsilon_t = \alpha + \gamma\varepsilon_{t-1} + ct + \sum_{k=1}^{K}\alpha_k \Delta\varepsilon_{t-k} + u_{3t}$$

- самое простое правило подбора K - включать 2 лага при объеме выборки меньше 81 наблюдения, 3 лага при объеме выборки от 81 до 256 наблюдений и т.д. [Канторович]

Альтернативные тесты

- PP-тест (тест Филлипса-Перрона): Используется при нарушении гипотезы о некоррелированности, гомоскедастичности и нормальности отклонений в тестируемой модели.
- PP-тест рекомендуется также к использованию в случаях наличия ярко выраженной сезонности и структурных сдвигов (скачков).

Подбор параметров модели ARMA

Идентификация параметров ARMA(p,q)

Автокорреляционная функция

Пусть X – некоторый временной ряд,

тогда его теоретическая АКФ (AC) имеет вид:

$$\rho(\tau) = \frac{1}{Var(X_t)} E\{(X_t - \mu)(X_{t-\tau} - \mu)\}$$

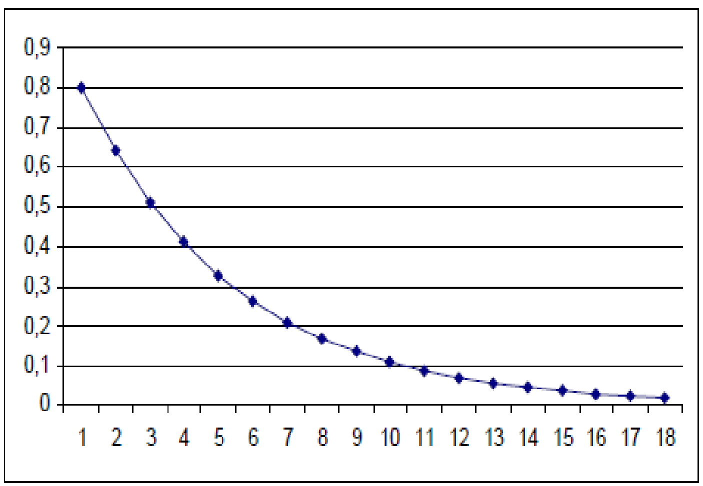


Рис. 1. Значения автокорреляционной функции процесса AR(1) при значении коэффициента равном 0,8

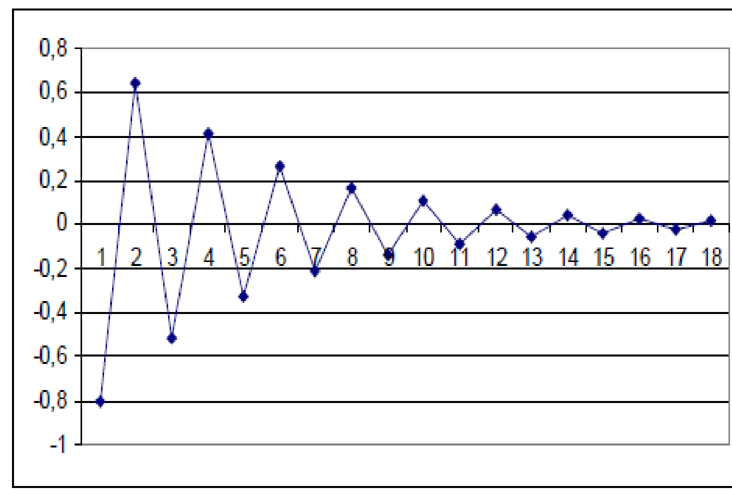


Рис. 2. Значения автокорреляционной функции процесса AR(1) при значении коэффициента равном −0,8

Идентификация параметров ARMA(p,q)

Частная автокорреляционная функция

Частная АКФ (PAC) определяется из системы

линейных уравнений Юла-Уокера,

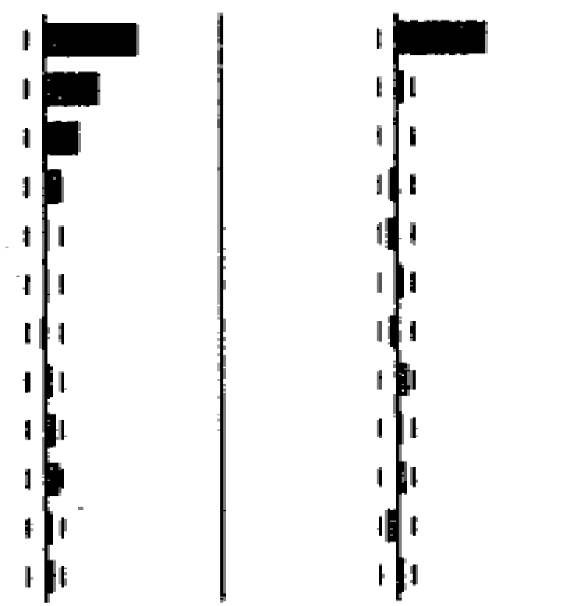
связывающей значения АКФ и

частной АКФ

$$
\begin{cases}
\rho_1 = \varphi_{k1}\cdot 1 + \varphi_{k2}\rho_1 + \varphi_{k3}\rho_2 + .. + \varphi_{kk}\rho_{k-1} \\
\rho_2 = \varphi_{k1}\rho_1 + \varphi_{k2}\cdot 1 + \varphi_{k3}\rho_1 + .. + \varphi_{kk}\rho_{k-2} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\rho_k = \varphi_{k1}\rho_{k-1} + \varphi_{k2}\rho_{k-2} + \varphi_{k3}\cdot\rho_{k-3} + .. + \varphi_{kk}\cdot 1.
\end{cases}
$$

Идентификация параметров ARMA

Вид коррелограмм AC и PAC для ρ>0

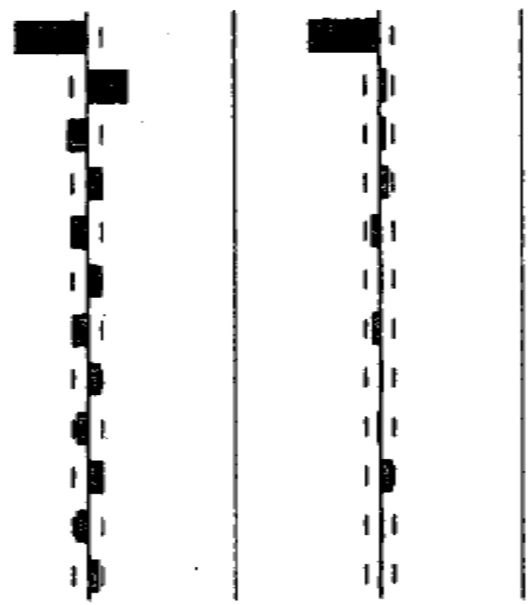


	AC	PAC	Q-Stat	Prob
1	0.539	0.539	116.40	0.000
2	0.319	0.041	157.37	0.000
3	0.190	0.004	171.91	0.000
4	0.092	-0.029	175.35	0.000
5	0.014	-0.044	175.43	0.000
6	0.012	0.033	175.50	0.000
7	-0.013	-0.026	175.56	0.000
8	0.025	0.059	175.81	0.000
9	0.042	0.018	176.52	0.000
10	0.069	0.042	178.47	0.000
11	0.027	-0.051	178.78	0.000
12	0.036	0.028	179.32	0.000

AR(1). $Y_t = 0.5Y_{t-1} + \varepsilon_t$

Идентификация параметров ARMA

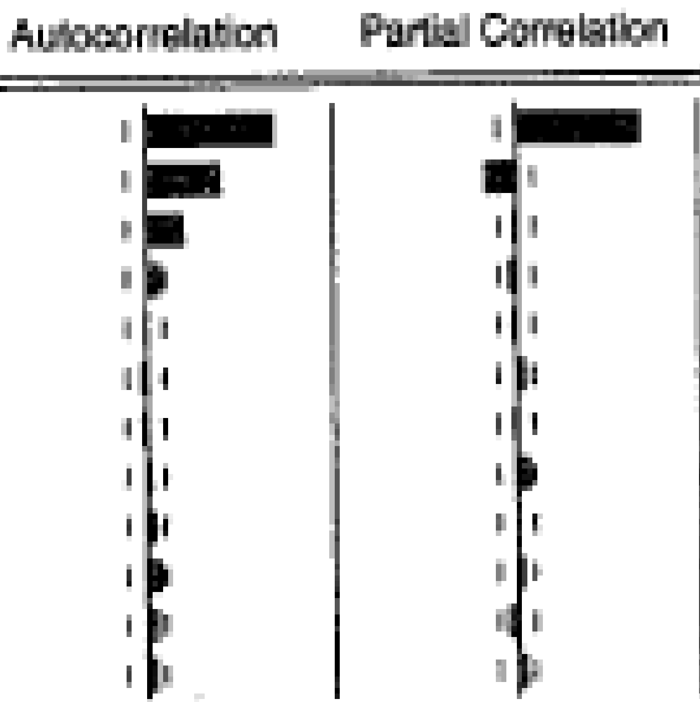
Вид коррелограмм AC и PAC для $\rho>0$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.500	-0.500	100.19	0.000
		2	0.281	0.041	131.88	0.000
		3	-0.125	0.041	138.15	0.000
		4	0.104	0.063	142.49	0.000
		5	-0.106	-0.049	147.01	0.000
		6	0.090	0.009	150.33	0.000
		7	-0.096	-0.043	154.11	0.000
		8	0.080	0.011	156.70	0.000
		9	-0.068	-0.010	158.57	0.000
		10	0.103	0.074	162.91	0.000
		11	-0.081	0.009	165.60	0.000
		12	0.063	-0.002	167.23	0.000

AR(1). $Y_t = -0.5Y_{t-1} + \varepsilon_t$

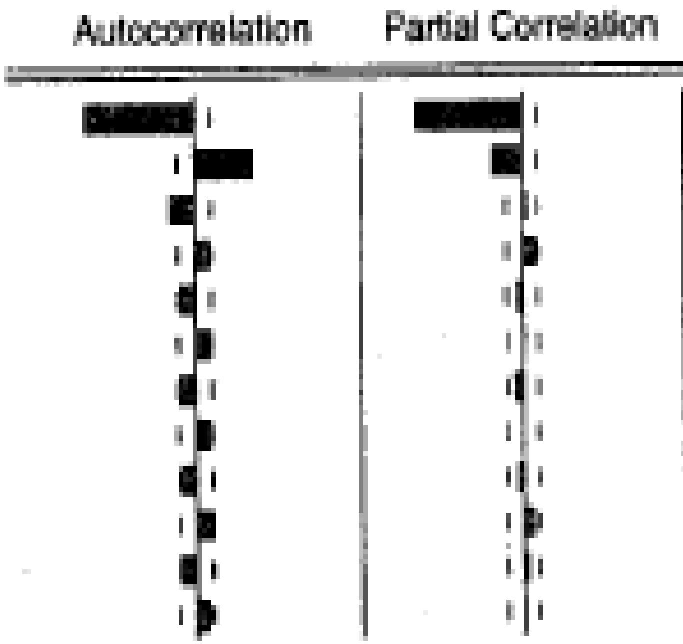
Идентификация параметров ARMA



	AC	PAC	Q-Stat	Prob
1	0.700	0.700	196.54	0.000
2	0.403	-0.171	261.80	0.000
3	0.203	-0.016	278.34	0.000
4	0.072	-0.037	280.46	0.000
5	-0.006	-0.023	280.47	0.000
6	-0.021	0.035	280.64	0.000
7	-0.022	-0.016	280.84	0.000
8	0.017	0.071	280.95	0.000
9	0.049	0.008	281.93	0.000
10	0.071	0.025	283.99	0.000
11	0.051	-0.043	285.05	0.000
12	0.048	0.045	286.00	0.000

AR(2). $Y_t = 0.8Y_{t-1} - 0.2Y_{t-2} + \varepsilon_t$.

Идентификация параметров ARMA

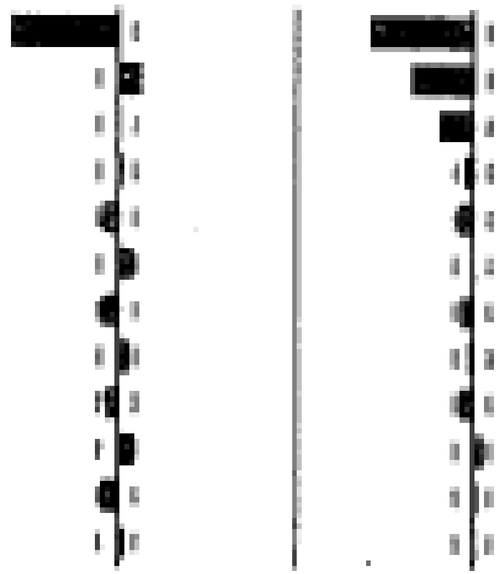


	AC	PAC	Q-Stat	Prob
1	-0.670	-0.670	179.75	0.000
2	0.353	-0.173	229.82	0.000
3	-0.147	0.028	238.48	0.000
4	0.087	0.083	241.55	0.000
5	-0.088	-0.032	244.67	0.000
6	0.090	0.009	247.99	0.000
7	-0.097	-0.042	251.78	0.000
8	0.088	0.007	254.96	0.000
9	-0.086	-0.030	257.98	0.000
10	0.106	0.062	262.57	0.000
11	-0.092	0.029	266.04	0.000
12	0.071	0.010	268.12	0.000

$$\mathrm{AR}(2).\quad Y_t = -0.8Y_{t-1} - 0.2Y_{t-2} + \varepsilon_t$$

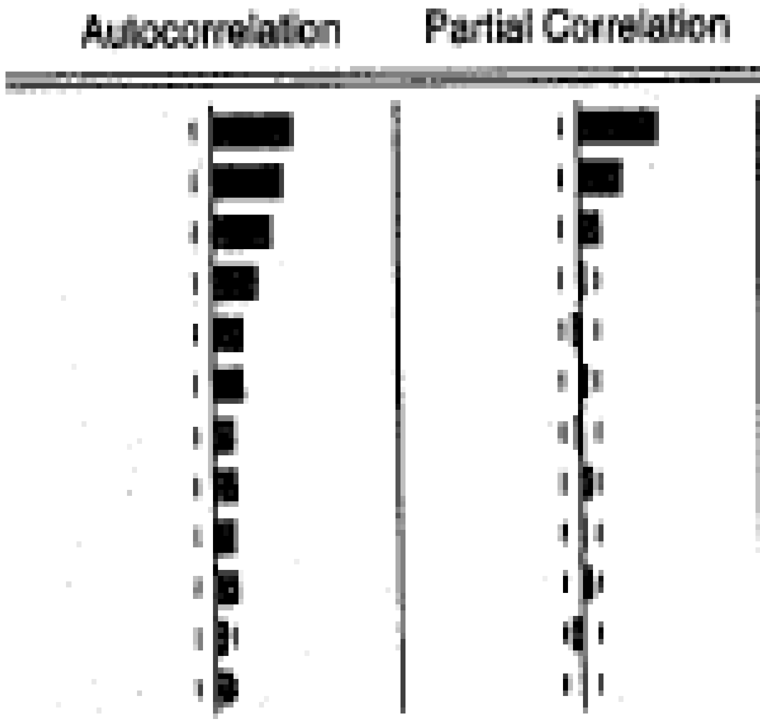
Идентификация параметров ARMA

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.593	-0.593	140.88	0.000
		2	0.124	-0.351	147.01	0.000
		3	0.004	-0.185	147.02	0.000
		4	0.026	-0.034	147.29	0.000
		5	-0.069	-0.068	149.21	0.000
		6	0.076	0.003	151.55	0.000
		7	-0.074	-0.050	153.79	0.000
		8	0.056	-0.014	155.06	0.000
		9	-0.055	-0.068	156.32	0.000
		10	0.088	0.050	159.47	0.000
		11	-0.077	0.024	161.89	0.000
		12	0.035	0.010	162.40	0.000



$$\mathrm{MA}(2).\quad Y_t = \varepsilon_t - 0.9\varepsilon_{t-1} + 0.2\varepsilon_{t-2}.$$

Идентификация параметров ARMA



	AC	PAC	Q-Stat	Prob
1	0.449	0.449	80.885	0.000
2	0.389	0.234	141.60	0.000
3	0.320	0.108	182.89	0.000
4	0.246	0.025	207.30	0.000
5	0.162	-0.037	217.92	0.000
6	0.161	0.040	228.44	0.000
7	0.100	-0.021	232.54	0.000
8	0.121	0.053	238.50	0.000
9	0.102	0.019	242.73	0.000
10	0.123	0.052	248.91	0.000
11	0.057	-0.053	250.23	0.000
12	0.067	0.002	252.10	0.000

ARMA(1,1). $Y_t = 0.8Y_{t-1} + \varepsilon_t - 0.5\varepsilon_{t-1}$

Метод максимального правдоподобия ММП

Метод максимального правдоподобия для КЛРМ

- Рассмотрим, как этот метод работает для классической регрессионной модели

$$Y_i = X_i'\beta + \varepsilon_i \qquad \varepsilon \sim N(0, \sigma_\varepsilon^2 I)$$

- Функция плотности Y_i при данных X_i имеет вид:

$$f\left(Y_i \mid \beta, \sigma_\varepsilon^2\right) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left\{-\frac{1}{2}\frac{\left(Y_i - X_i'\beta\right)^2}{\sigma_\varepsilon^2}\right\}$$

Оценка ММП для классической линейной регрессии

- Соответствующая логарифмическая функция правдоподобия будет иметь вид:

$$\ln L\left(\beta, \sigma_\varepsilon^2\right) = \sum_{i=1}^{n} \ln L_i\left(\beta, \sigma_\varepsilon^2\right) = \sum_{i=1}^{n} \ln f\left(Y_i \mid \beta, \sigma_\varepsilon^2\right) =$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^{n}\left(Y_i - X_i'\beta\right)^2$$

- Далее находятся оценки параметров, обеспечивающие максимум этой функции

Оценка ММП для классической линейной регрессии

- Условия первого порядка

$$\frac{\partial \ln L\left(\beta, \sigma_\varepsilon^2\right)}{\partial \beta} = 0 \qquad \frac{\partial \ln L\left(\beta, \sigma_\varepsilon^2\right)}{\partial \sigma_\varepsilon^2} = 0$$

- приводят к уравнениям

$$\frac{1}{\sigma_\varepsilon^2} \frac{\partial}{\partial \beta} \sum_{i=1}^{n} X_i \left(Y_i - X_i'\beta\right) = 0$$

$$-\frac{n}{2\sigma_\varepsilon^2} + \frac{1}{2\sigma_\varepsilon^4} \sum_{i=1}^{n} \left(Y_i - X_i'\beta\right)^2 = 0$$

Оценка ММП моделей ARMA

Пусть y_t подчиняется процессу ARMA(1,1)

Введем обозначения:

$$y_1^* = y_1, \quad y_2^* = y_2 + \theta_1 y_1, \quad \ldots \quad y_t^* = y_t + \theta_1 y_{t-1} + \ldots + \theta_1^{t-1} y_1$$

И будем оценивать модель вида

$$y_t^* = \delta^* + \phi_1 y_{t-1}^* + \varepsilon_t$$

$$\varepsilon_t \sim iid N(0, \sigma^2)$$

Оценка ММП моделей ARMA

Логарифм функции правдоподобия для ARMA(1,1) будет иметь вид:

$$l(\delta, \phi_1, \theta_1, \sigma^2) =$$

$$= \text{const} - \frac{n}{2}\ln\sigma^2 + \frac{1}{2}\ln(1-\phi_1^2) - \frac{1-\phi_1^2}{2\sigma^2}\left(y_1^* - \frac{\delta^*}{1-\phi_1}\right)^2$$

$$- \frac{1}{2\sigma^2}\sum_{t=2}^{n}(y_t^* - \delta^* - \phi_1 y_{t-1}^*)^2.$$

Критерии качества подгонки

Информационный критерий Акаике (AIC)

$$AIC = -2\frac{l}{T} + 2\frac{k}{T}$$

Информационный критерий Шварца (BIC)

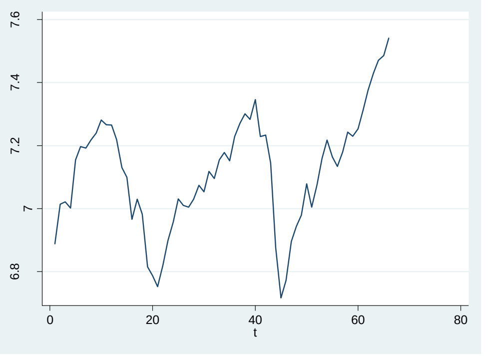
$$BIC = SC = -2\frac{l}{T} + \frac{k\log T}{T}$$

l – логарифм функции правдоподобия,

к – число оцениваемых параметров.

Чем ниже значения критериев, тем лучше результат

ARIMA (1,0,1) regression



Sample: 1 - 66

Number of obs = 66
Wald chi2(2) = 215.46
Prob > chi2 = 0.0000

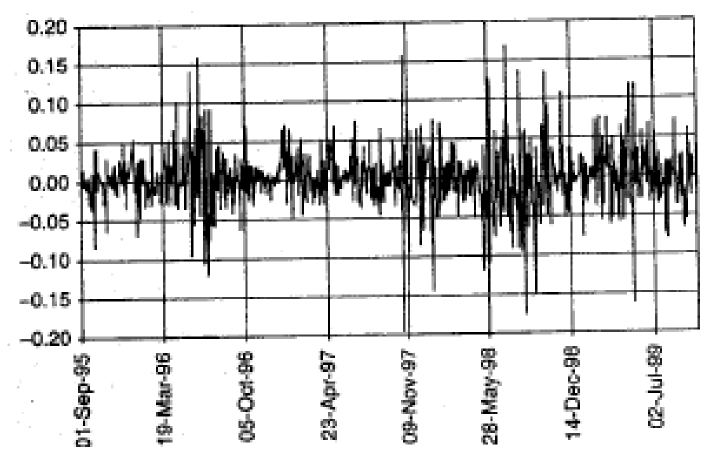
Log likelihood = 81.4261

sp500	Coef.	Std. Err.	z	P>z	[95% Conf.Interval]	
_cons	7.135755	.1141207	62.53	0.000	6.912083	7.359427
ARMA						
L1.ar	.9018984	.0822384	10.97	0.000	.7407141	1.063083
L1.ma	.3435668	.1170013	2.94	0.003	.1142484	.5728852
/sigma	.0692228	.0065627	10.55	0.000	.05636	.0820855

GARCH модели ВР

Из эмпирических наблюдений над поведением ВР процентных ставок, валютных курсов и т.п. было замечено, что наблюдения с большими отклонениями от среднего и с малыми отклонениями склонны к образованию кластеров

Однодневные приращения индекса РТС



GARCH модели ВР

- Это явление оказалось удобно моделировать зависимостью дисперсии ошибок от предыстории

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2.$$

- Простейшая модель этого класса ARCH(1) запишется в виде:

$$y_t = x_t' \beta + u_t.$$

$$u_t = \varepsilon_t \left(\alpha_0 + \alpha_1 u_{t-1}^2\right)^{1/2}, \quad \varepsilon_t \sim iidN(0,1).$$

GARCH модели ВР

- Если сформулировать зависимость дисперсии от предыстории в более общем виде

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_q \sigma_{t-q}^2$$

- то это приведет к моделям GARCH(p,q).
- Простейшую модель этого вида GARCH(1,1) можно записать так:

$$y_t = x_t'\beta + u_t$$

$$u_t = \varepsilon_t \left(\alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 \sigma_{t-1}^2\right)^{1/2},$$

$$\varepsilon_t \sim iidN(0,1)$$

ARCH family regression

- Sample: 1 - 66 Number of obs = 66
- Distribution: Gaussian Wald chi2(.) = .
- Log likelihood = 33.74703 Prob > chi2 = .
- --
- sp500 | Coef. Std. Err. z P>|z| [95% Conf. Interval]
- -------------+--
- _cons | 7.071681 .0140875 501.98 0.000 7.04407 7.099292
- -------------+--
- ARCH
- L1. arch | 1.004295 .6102791 1.65 0.100 -.1918305 2.20042

  L1. garch | -.0887646 .2598499 -0.34 0.733 -.5980609 .4205318
- _cons | .0053252 .0055901 0.95 0.341 -.0056312 .0162816
- --

Метод оценивания моделей GARCH

- Модели GARCH оцениваются методом максимального квазиправдоподобия, что означает следующее:
  - используется гипотеза о нормальности, даже если ошибка ненормальна,
  - делается специальная корректировка при вычислении стандартных ошибок

Условия состоятельности оценок

- 1. Условие верной идентификации первых моментов

$$E\left(\varepsilon_t \mid z_{t-1}\right)=0, \quad E\left(\varepsilon_t^2 \mid z_{t-1}\right)=1$$

- 2. Условие стационарности

$$E\left\{\ln\left(\alpha_1\varepsilon_t^2\right) \mid z_{t-1}\right\}<0 \quad \alpha_1>0$$

- 3. Условие асимптотической нормальности

$$E\left(\varepsilon_t^4 \mid z_{t-1}\right)-\textit{ограничено}$$

Достоинства моделей GARCH

- Метод позволяет оценивать регрессии

$$y_t = x_t'\beta + u_t$$

- с не гауссовскими (не нормальными) распределениями ошибок при наличии тяжелых хвостов,
- успешно справляется с сериальной корреляцией квадратов ошибок
- несложно приспосабливаются для моделирования финансовых ВР

Динамические регрессионные модели

Построение зависимостей между различными ВР

- Для ВР появляется возможность строить более сложные и реалистичные модели явлений, учитывающие запаздывание реакции зависимой переменной на изменения независимых переменных, а также память в самой зависимой переменной например:

$$Y_t = a_0 + b_0 X_t + b_1 X_{t-1} + \gamma Y_{t-1} + \varepsilon_t$$

Построение зависимостей между различными ВР

- Если ряды Xt и Yt не являются стационарными, зависимость между ними может оказаться ложной.
- Надо обязательно выделять тренд и сезонность.
- Признаками ложной регрессии являются высокий R^2 при низкой статистике Дарбина-Уотсона DW.

Например

$$y_t = \underset{(-5.77)}{-2.79} - \underset{(-21.5)}{0.52}\, x_t; \quad R^2 = 0.607, \quad DW = 0.057$$

Тест Дарбина-Уотсона

- **Тест Дарбина-Уотсона** используется для проверки гипотезы об отсутствии автокорреляции первого порядка в остатках регрессионной модели.
- Пусть (e_1,…,e_n) - это вектор остатков линейной регрессии по k независимым переменным.
- Предполагая, что остатки образуют процесс AR(1) $e_t = \rho e_{t-1} + \upsilon_t$,

где υ_t – последовательность независимых нормальных случайных величин.

Тест Дарбина-Уотсона

- Тест основан на проверке гипотезы об отсутствии автокорреляции

$$H_o: \rho = 0,$$

- критерием служит статистика Дарбина-Уотсона, которая рассчитывается по следующей формуле:

$$DW = \sum_{t=2} (e_t - e_{t-1})^2 \Big/ \sum_{t=2} e_t^2$$

$$DW \approx 2(1-\rho)$$

Тест Дарбина-Уотсона

- Поскольку $DW \approx 2(1 - \rho)$, где ρ - коэффициент корреляции между e_i и e_{i-1}, значения DW находятся в промежутке от 0 до 4.
- При $\rho = 0$ DW близка к 2.
- Близость DW к нулю говорит о положительной автокорреляции, к 4 - об отрицательной.
- На практике проверка гипотезы H_0 об отсутствии автокорреляции остатков осуществляется с помощью сравнения статистики DW с теоретическими значениями d_l и d_u для заданного числа наблюдений n, числа независимых переменных модели k и уровня значимости α

Тест Дарбина-Уотсона

- $0 < DW < d_l$ - гипотеза H_0 отвергается, есть положительная автокорреляция;
- $d_l < DW < d_u$ - зона неопределённости;
- $d_u < DW < 4 - d_u$ - гипотеза H_0 не отвергается, автокорреляции нет;
- $4 - d_u < DW < 4 - d_l$ - зона неопределённости;
- $4 - d_l < DW < 4$ - гипотеза H_0 отвергается, есть отрицательная автокорреляция.

Проблемы моделирования зависимостей между ВР

- Временные ряды Xt и Yt – имеют различный тип нестационарности.
- Например, ряд Yt является нестационарным и интегрированным порядка k , а Xt является TS-рядом.
- В таких случаях может иметь место ложная регрессионная зависимость, а операции исключения тренда или перехода к разностям не гарантируют принадлежности случайных отклонений классу «гауссов белый шум»

Проблемы моделирования зависимостей между ВР

- Временные ряды Xt   и Yt имеют различный порядок интегрирования, т.е. Xt ~ I(k), Yt~ I(l) t , k не равно l
- при этом ряд остатков модели содержит стохастический тренд.
- В таком случае может иметь место ложная регрессионная зависимость
- Целесообразно строить модель на основе временных рядов разностей соответствующих порядков

Коинтеграция ВР

- Не ложная регрессия между нестационарными ВР возможна, если ВР являются коинтегрированными.
- Это означает либо стационарность ошибки для некоторой линейной комбинации

$$\alpha\, y_t + x_t'\beta = u_t$$

- Либо в более общем случае, линейная комбинация рядов с порядком интегрирования d имеет более низкий порядок интегрирования d-1

Построение зависимостей между различными ВР

- Серьезной проблемой регрессионных зависимостей для ВР является автокорреляция ошибок.
- Иногда выбор удачной динамической спецификации, формы тренда или учет сезонности позволяют ее избежать.
- Отдельный сложный вопрос – направление причинно-следственной связи.

Причинность по Грэнджеру

- Способ выяснить статистическую причинность предлагает тест Грэнджера
- Оцениваются две регрессии:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \beta_1 x_{t-1} + \ldots + \beta_p x_{t-p} + \varepsilon_t$$
$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \ldots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t$$

- И для каждой проверяется гипотеза

$$\beta_1 = \ldots = \beta_p = 0$$

Популярные спецификации регрессионных моделей для ВР

- Модель частичного приспособления
- Модель адаптивных ожиданий
- Модель коррекции ошибок
- Модель векторной авторегрессии (VAR)

Модель частичного приспособления

<u>Пример.</u> Рассмотрим зависимость между оптимальным (ненаблюдаемым) потреблением бензина и ценами на нефть:

$$Y_t^* = a + bX_t + \varepsilon_t$$

Реальное потребление постепенно приближается к оптимальному по правилу

$$Y_t - Y_{t-1} = (1-\lambda)\left(Y_t^* - Y_{t-1}\right)$$

Итоговая модель

$$Y_t = (1-\lambda)\alpha + (1-\lambda)\beta X_t + \lambda Y_{t-1} + u_t,\ u \sim N\left(0, \sigma^2 I\right)$$

Модель адаптивных ожиданий

Пример. Рассмотрим зависимость между выпуском и оптимальным (ненаблюдаемым) объемом продаж

$$Y_t = a + bX_t^* + \varepsilon_t$$

Реальные продажи постепенно приближаются к оптимальным по правилу

$$X_t^* - X_{t-1}^* = (1-\lambda)\left(X_{t-1} - X_{t-1}^*\right)$$

Итоговая модель

$$Y_t = (1-\lambda)\alpha + (1-\lambda)\beta X_t + \lambda Y_{t-1} + u_t,\ u \sim MA(1)$$

Модель коррекции ошибок

Пример. Рассмотрим зависимость между продажами (Y) и затратами на рекламу (X)

$$Y_t = a_0 + b_0 X_t + b_1 X_{t-1} + \gamma Y_{t-1} + \varepsilon_t$$

Это уравнение часто переписывают в виде

$$\Delta Y_t = b_1 \Delta X_t - (1-\gamma)(Y_{t-1} - a - bX_{t-1}) + \varepsilon_t$$

где $a = a_0/(1-\gamma),\ b = (b_0 + b_1)/(1-\gamma)$

Выражение в скобках $Y_{t-1} - a - bX_{t-1} = u_t$ представляет собой «остаток равновесия»,

(1-γ) – скорость коррекции

Векторная авторегрессия (VAR)

- VAR это система одновременных уравнений, которая состоит из одномерных моделей ARMA

$$Y_t = a_1 + b_{11}Y_{t-1} + b_{12}X_{t-1} + \varepsilon_{1t}$$

$$X_t = a_2 + b_{21}Y_{t-1} + b_{22}X_{t-1} + \varepsilon_{2t}$$

- $\varepsilon_1, \varepsilon_2$ - белые шумы, которые могут быть коррелированы

Преимущества VAR

- Модель может быть более экономной, включая меньше лагов
- Прогноз может быть точнее
- В модели не нужно уметь различать зависимые и независимые переменные
- Модель может быть оценена обычным МНК и оценки будут состоятельными, поскольку белый шум предполагается независимым от истории

Замечания о качестве эконометрических прогнозов

- При **неизменных внешних условиях**, когда эконометрическая модель и механизм порождения данных соответствуют друг другу прогноз, вычисленный как условное ожидание, будет оптимальным, т.е. несмещенным и эффективным.
- Различия прогноза и дальнейшей реализации процесса будут обусловлены только ошибкой, которую невозможно точно предсказать.
- Однако неизменность внешних условий в экономике обеспечить затруднительно, и в этом причина частой несостоятельности прогнозов.

Замечания и качестве эконометрических прогнозов

- Теория прогнозирования, основанная на предположениях о стационарности процессов и постоянстве параметров моделей является неадекватной.
- С этими предположениями связана и гипотеза о нормальности, поскольку в нормальном законе распределения вероятностей дисперсия и математическое ожидание предполагаются неизменными.
- Однако использование моделей типа GARCH или стохастической волатильности позволяет отказаться от нереалистичных предположений и улучшить качество прогноза.

Спасибо за внимание!