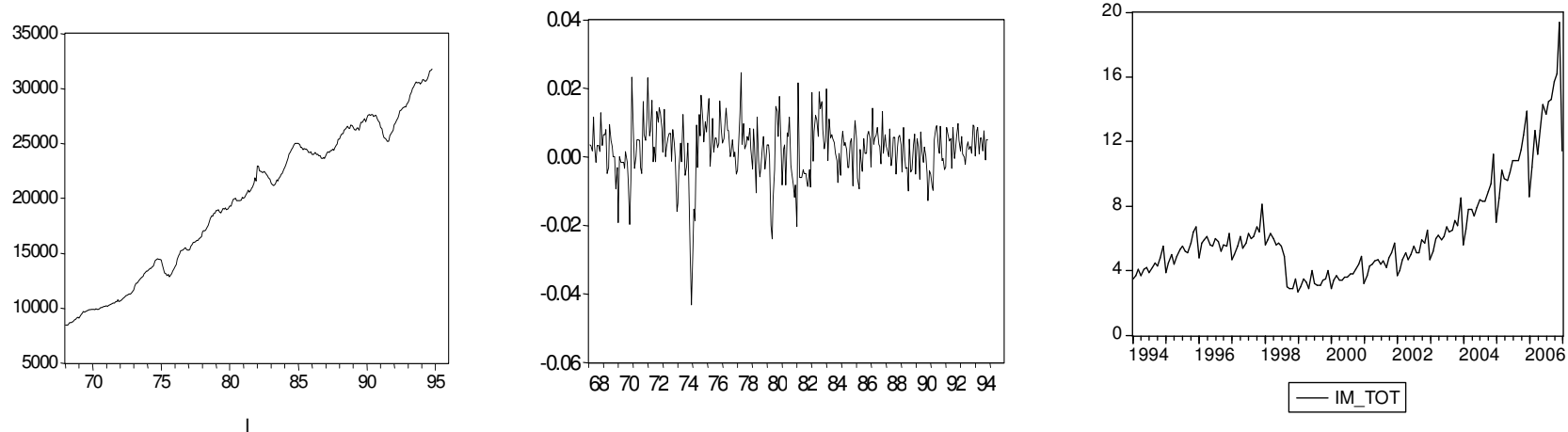


Анализ временных рядов и прогнозирование

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 различных временных рядов



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

В рамках нашего курса мы рассмотрим:

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

Цели анализа ВР

- краткое описание характерных особенностей ряда;
- подбор статистической модели, описывающей ВР;
- предсказание будущих значений на основе прошлых наблюдений;
- управление процессом, порождающим ВР.

Стадии анализа ВР

- графическое представление и описание поведения ВР;
- выделение и удаление закономерных составляющих ВР, зависящих от времени: тренда, сезонных и циклических составляющих;
- выделение и удаление низко- или высокочастотных составляющих процесса (фильтрация);
- исследование случайной составляющей ВР, оставшейся после удаления перечисленных выше составляющих;
- построение математической модели для описания случайной составляющей и проверка ее адекватности;
- прогнозирование будущего развития процесса, представленного ВР;
- исследование взаимодействий между различными ВР.

Методы анализа ВР

- корреляционный анализ позволяет выявить существенные периодические зависимости и их лаги (задержки) внутри одного процесса (автокорреляция) или между несколькими процессами (кросскорреляция);
- спектральный анализ позволяет находить периодические и квазипериодические составляющие ВР;
- сглаживание и фильтрация предназначены для преобразования ВР с целью удаления из них высокочастотных или сезонных колебаний;
- модели авторегрессии и скользящего среднего оказываются особенно полезными для описания и прогнозирования процессов, проявляющих однородные колебания вокруг среднего значения;
- прогнозирование позволяет на основе подобранной модели поведения ВР предсказывать его значения в будущем.

При анализе ВР принято выделять 4 компоненты:

- тренд (Т) – плавно изменяющаяся компонента, описывающая чистое влияние долговременных факторов (рост населения, изменение структуры возрастного состава и т.д.);
- 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 (S) – состоит из последовательности почти повторяющихся циклов (объем продаж накануне Нового Года, объем перевозок пассажиров городским транспортом);
- случайная компонента (e) – остается после полного вычленения закономерных компонент.

ВР представляет собой

- либо сумму этих компонент $X=T+C+S+e$ в аддитивной модели,
- либо произведение $X=T*C*S*e$ 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 - $T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_k t^k$
- логарифмическая - $T_t = \exp(\beta_0 + \beta_1 t)$

Представление о характере тренда можно получить из графика ВР.

Выделение циклической компоненты

Чтобы выделить циклическую компоненту, работают с ВР, не содержащими сезонности (например, годовыми ВР).

Из мультипликативной модели $X = T \cdot C \cdot S \cdot e$

C_t выделяется следующим образом:

- Удаляют тренд из модели $(C \cdot \varepsilon)_t = \frac{X_t}{T_t}$
- Вычисляют процентную долю, приходящуюся на $(C \cdot \varepsilon)_t : (C \cdot \varepsilon)_t \cdot 100\%$.
- Вычисляют скользящее среднее для процентной доли:
 $MA[(C \cdot \varepsilon)_t \cdot 100\%]$
эту величину и считают циклической компонентой.
- Для построения прогноза бывает необходимо выделять и исследовать случайную компоненту :
 $[(C \cdot \varepsilon)_t \cdot 100\%] / MA[(C \cdot \varepsilon)_t \cdot 100\%]$.

По полученному графику можно определить точки поворота цикла.

Выделение сезонной компоненты

Для выделения сезонной компоненты
поступают следующим образом:

- сглаживают ВР, т.е. вычисляют

$$MA[X_t] = (T \cdot C)_t$$

- вычисляют компоненту

$$(S \cdot \varepsilon)_t = \frac{X_t}{(T \cdot C)_t} = \frac{X_t}{MA(X_t)}$$

- изолируют сезонную компоненту
усреднением $(S \cdot \varepsilon)_t$

Рассмотрим пример выполнения п.3 для данных о количестве миль, сделанных за квартал автобусом

Год	1-ый кв.	2-ой кв.	3-ий кв.	4-ый кв.
1986			180.21	71.17
1987	45.19	104.61	179.86	72.89
1988	46.80	99.82	177.24	77.06
1989	43.33	103.63	175.44	74.96
1990	44.78	104.02	178.17	70.14
1991	49.88	101.24		
медиана	45.19	103.63	178.17	72.89

В качестве сезонного индекса выбирается медиана,
т.к. она не реагирует на выбросы.

По индексам легко определить значимость влияния
отдельного сезона.

Знание сезонных индексов позволяет ответить на вопросы типа нижеследующих:

- Является ли настоящее увеличение безработицы естественным для данного периода времени?
- Не маскирует ли предрождественский наплыв покупателей переход к новому периоду экономического цикла?
- Если в будущем году предполагается продать 400 единиц товара, то каковы будут предполагаемые ежемесячные продажи с учетом сезонности?

Для ответа на подобные вопросы необходимо:

- Скорректировать влияние сезонной компоненты, чтобы более четко обнаружились основные тенденции;
- Инкорпорировать влияние сезонности в прогноз.

Для достижения первой цели проводится десезонализация:

$$(T \cdot C \cdot \varepsilon)_t = \frac{X_t}{S_t}$$

где под S_t понимается соответствующий индекс

Пример. Исследование ВР числа транзакций универсама

№ квартал а	Число транзакци й (в тыс.)	Сезонный индекс	Число транзакций, скорр. на сезонность $(T \cdot C \cdot \varepsilon)_t = X_t / S_t$	Годовая норма транзакций $4 \cdot X_t / S_t$
1	17.36	70	24.80	99.20
2	21.11	85	24.84	99.36
3	21.68	92	23.57	94.28
4	32.35	153	21.08	84.32

Из таблицы видно, что несмотря на большое число транзакций в 4-ом квартале, годовая норма транзакций после корректировки на сезонность падает к 4-ому кварталу довольно заметно.
Возможно – это начало циклического спада

Пример. Прогнозирование объема продаж

Пусть должно быть продано 200 единиц товара. Что можно ожидать для поквартальных объемов продаж?

№ квартала	X без S (1)	Сезонный индекс S (2)	X с S : (1)*(2)/100
1	50	123	61.5
2	50	108	54.0
3	50	79	39.5
4	50	90	45.0

Замечание

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

$$X_t = \beta_0 + \beta_1 t + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \varepsilon_t$$

Построение прогноза для ВР

- Задача прогнозирования состоит в том, чтобы по имеющимся наблюдениям ВР предсказать неизвестные будущие значения
- Прогнозирование в бизнесе играет очень большую роль, поскольку оно 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.

Подход Бокса-Дженкинса ARMA

- Методы ARMA предназначены для прогнозирования либо стационарных ВР, либо для ВР, которые могут быть преобразованы к стационарным.
- К таким рядам относятся, например, остатки регрессионных моделей типа

$$X_t = \beta_0 + \beta_1 t + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \varepsilon_t$$

Стационарность

- Ряд является *стационарным*, если он совершает колебания вокруг своего математического ожидания .
- Сами значения ряда не являются, как правило, независимыми, но корреляция между членами ряда зависит только от расстояния S между ними.

Подход Бокса-Дженкинса ARMA

- Предположение о том, что существует связь между соседними значениями ВР и составляет основу методов ARMA.
- Именно эта гипотеза позволяет предсказать значения $\varepsilon_{t+1}, \varepsilon_{t+2}, \dots$ на основании известных значений $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t$
- затем на основании регрессионной модели
$$X_t = \beta_0 + \beta_1 t + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \varepsilon_t$$
- строятся будущие значения X_{t+1}, X_{t+2}, \dots

Методология Бокса-Дженкинса

- Ряд ε_t считается реализацией случайного процесса ARMA(p,q), и в зависимости от типа этого процесса строится прогноз.

- Например, если $\varepsilon_t \sim \text{AR}(1)$, то

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

- где $E(u_t) = 0$, $D(u_t) = \sigma^2$.

- Тогда прогнозное значение

$$\hat{\varepsilon}_{t+1} = \hat{a}_0 + \hat{a}_1 \varepsilon_t$$

Методология Бокса-Дженкинса

- Если $\varepsilon_t \sim \text{MA}(1)$, то $\varepsilon_t = a_0 + b_1 u_{t-1} + u_t$
- Если $\varepsilon_t \sim \text{ARMA}(1,1)$, то

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

- и прогнозное значение

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

- Если исходный ряд не является стационарным, то иногда его можно сделать стационарным, перейдя к первым (или вторым) разностям.

Алгоритм анализа 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 АС и РАС.
4. Оценивание параметров методом максимального правдоподобия и выбор наилучшей модели (критерии Акаике, Шварца).
5. Диагностическая проверка (анализ коррелограмм АС и РАС).
6. Прогнозирование.

Тестирование стационарности ВР (тест Дики-Фуллера)

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

$$y_t = b_1 y_{t-1} + \epsilon_{1t},$$

$$y_t = a_2 + b_2 y_{t-1} + \epsilon_{2t},$$

$$y_t = a_2 + b_3 y_{t-1} + c_3 t + \epsilon_{3t}.$$

- Оценивается уравнение $\Delta y_t = \gamma y_{t-1} + \epsilon_t$
- И проверяется гипотеза $H_0: \gamma = 0$
- Которая соответствует наличию единичного корня, т.е. нестационарности ВР

Автокорреляционная функция

Пусть X – некоторый временной ряд,
тогда его теоретическая АКФ имеет вид:

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

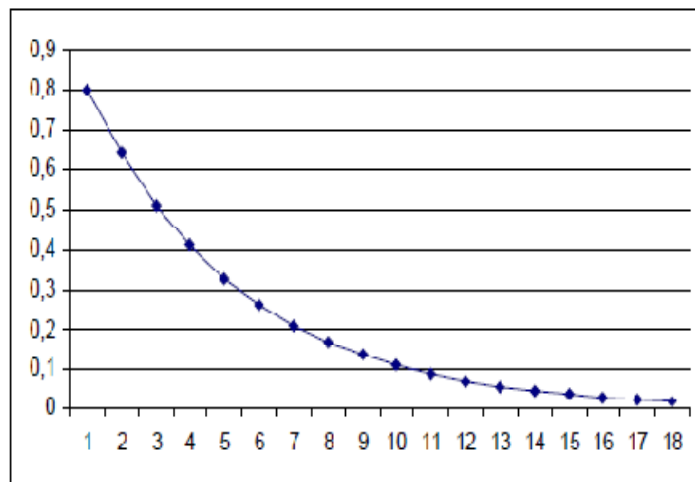


Рис. 1. Значения автокорреляционной функции процесса AR(1) при значении коэффициента равном 0,8

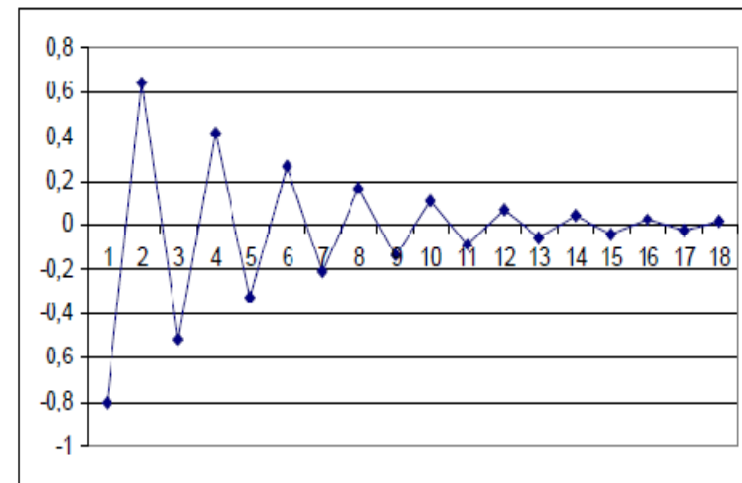


Рис. 2. Значения автокорреляционной функции процесса AR(1) при значении коэффициента равном -0,8

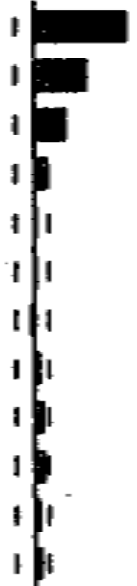

Частная автокорреляционная функция

Частная АКФ определяется из системы линейных уравнений Юла-Уокера, связывающей значения АКФ и частной АКФ

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

Идентификация параметров ARMA



Вид коррелограмм АС и РАС для $\rho > 0$

Autocorrelation	Partial Correlation	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

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


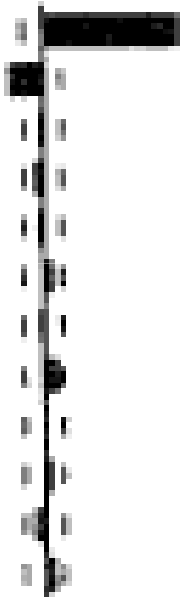
Идентификация параметров ARMA

Вид коррелограмм АС и РАС 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

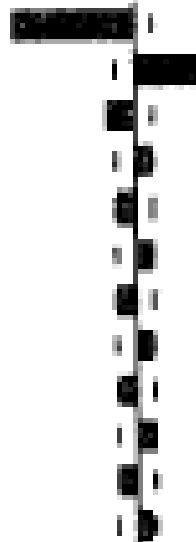
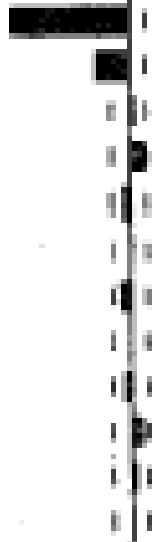
AR(1). $Y_t = -0.5Y_{t-1} + \epsilon_t$

Идентификация параметров ARMA

Autocorrelation	Partial Correlation	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	0.006	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

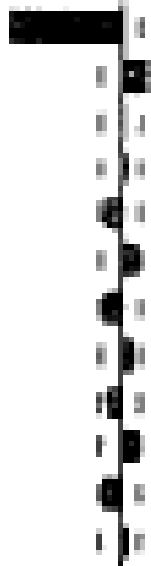
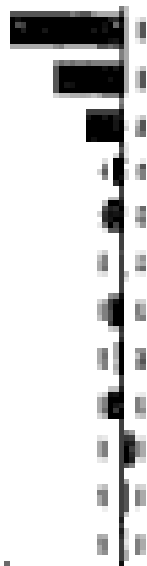
AR(2). $Y_t = 0.8Y_{t-1} - 0.2Y_{t-2} + \varepsilon_t$

Идентификация параметров ARMA

Autocorrelation	Partial Correlation	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	0.080	0.009	247.99	0.000
		7	-0.087	-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	-0.082	0.029	266.04	0.000
		12	0.071	0.010	268.12	0.000


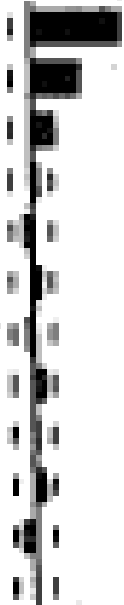
AR(2). $Y_t = -0.8Y_{t-1} - 0.2Y_{t-2} + \varepsilon_t$

Идентификация параметров ARMA

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.533	-0.533	140.88	0.000
		2	0.124	-0.351	147.01	0.000
		3	0.034	-0.185	147.02	0.000
		4	0.026	-0.004	147.29	0.000
		5	-0.059	-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	-0.058	156.32	0.000
		10	0.038	0.050	159.47	0.000
		11	-0.077	0.024	161.89	0.000
		12	0.035	0.010	162.40	0.000

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

Идентификация параметров ARMA

Autocorrelation	Partial Correlation		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	0.087	0.002	252.10	0.000

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

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,

k – число оцениваемых параметров.

Чем ниже значения критериев, тем лучше
результат

Оценивание параметров ВР

Пусть y_t подчиняется процессу ARMA(1,1)

Введем обозначения:

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

И будем оценивать модель вида

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

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

Оценивание параметров ВР

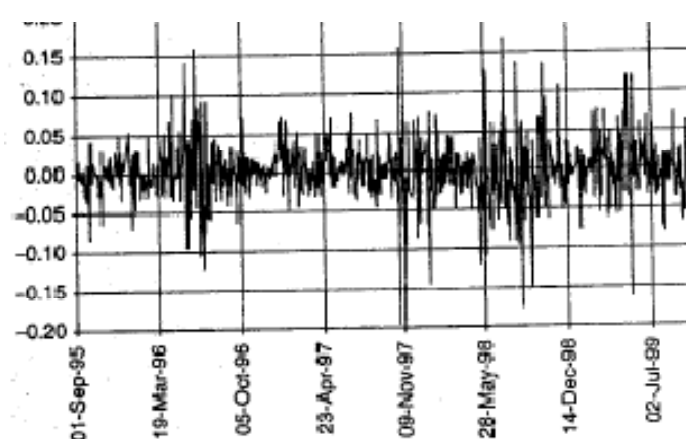
- Логарифм функции правдоподобия будет для ARMA(1,1) иметь вид:

$$\begin{aligned} l(\delta, \phi_1, \theta_1, \sigma^2) = \\ = \text{const} - \frac{n}{2} \ln \sigma^2 + \frac{1}{2} \ln(1 - \phi_1^2) - \frac{1 - \phi_1^2}{2\sigma^2} \left(y_1^* - \frac{\delta^*}{1 - \phi_1} \right)^2 \\ - \frac{1}{2\sigma^2} \sum_{t=2}^n (y_t^* - \delta^* - \phi_1 y_{t-1}^*)^2. \end{aligned}$$

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 BP

- Это явление оказалось удобно моделировать зависимостью дисперсии ошибок от предыстории

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

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

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

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

GARCH модели BP

- Если сформулировать зависимость дисперсии от предыстории в более общем виде

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

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

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

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

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

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

$$E(\varepsilon_t | z_{t-1}) = 0, E(\varepsilon_t^2 | z_{t-1}) = 1$$

- 2. Условие стационарности

$$E\{\ln(\alpha_1 \varepsilon_t^2) | z_{t-1}\} < 0$$

- 3. Условие асимптотической нормальности

$$E(\varepsilon_t^4 | z_{t-1}) - \text{ограничено}$$

$$\alpha_0, \alpha_1 > 0$$

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 приспособиваются для моделирования финансовых ВР

Построение зависимостей между различными ВР

- Для ВР появляется возможность строить более сложные и реалистичные модели явлений, учитывающие возможность запаздывающих влияний зависимой и независимой переменных, 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 y_t и x_t не являются стационарными, зависимость между ними может оказаться ложной.
 - Признаками ложной зависимости являются высокий R^2 при низкой статистике Дарбина-Уотсона, например
- $$y_t = -2.79 - 0.52 x_t; \quad R^2 = 0.607, \quad DW = 0.057$$
- $(-5.17) \quad (-21.5)$
- Надо обязательно выделять тренд и сезонность.

Коинтеграция ВР

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

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

Построение зависимостей между различными ВР

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

Причинность по Грэнджеру

- Способ выяснить статистическую причинность предлагает тест Грэнджера

- Оцениваются две регрессии:

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

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

- И для каждой проверяется гипотеза

$$\beta_1 = \dots = \beta_p = 0$$

Популярные спецификации регрессионных моделей для ВР

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

Модель частичного приспособления

Пример. 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

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

Итоговая модель

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

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

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

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 и затратами на рекламу

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

- 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 $a = a_0 / (1 - \gamma)$, $b = (b_0 + b_1) / (1 - \gamma)$
- член $Y_{t-1} - a - bX_{t-1} = u_t$ представляет собой «остаток равновесия»,
- $(1 - \gamma)$ – 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

Качество 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 история хранит множество случаев несостоятельности прогнозов.

Качество эконометрических прогнозов

- Теория прогнозирования, основанная на предположениях о стационарности процессов и постоянстве параметров (которые должны ухватываться моделью) 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.

Заключение-1

- Подводя итоги нашего курса и размышляя о целесообразности изучения эконометрических методов для работы на финансовых рынках, уместно процитировать высказывание Грэнджера и Тиммермана:
- «Для получения количественной оценки ожидаемой экономической прибыли и степени рыночной эффективности необходим совместный анализ информационных множеств и моделей, включающих:

Заключение-2

- информационное множество (Ω) публично доступной информации с биржевыми ценами и доступными результатами биржевой работы фондов и частных инвесторов;
- модель равновесных цен и модель премии за риск (ER)
- набор моделей предсказания (M), доступных в любой момент времени, включая метод их оценки
- технологии поиска (S) для выбора лучшей модели прогнозирования из набора моделей M

Заключение-3

- модель доступности информации (асимметрии), включая публичные версии частной информации, стоимости такой информации и затрат на ее преобразование в модели принятия решения M и S во времени (A);
- модель изменения во времени стоимости транзакций, ликвидности и доступных технологий торговли (L);
- модель истощения (старения) (GO) моделей предсказания M (моделей рыночной эффективности)»

Заключение-4

- Тогда можно определить два способа получения (измерения) экономической прибыли:
- практический (на основании Ω, ER, M, S)
- модельный (на основании $\Omega, ER, M, S, A, L, GO$) с многократным повторением вероятностных бутстрап-испытаний устойчивости модельных параметров