Содержание

1 Общие понятия **2**

2 Процесс белого шума **5**

3 Процесс авторегрессии **8**

4 Пример **19**

1 Общие понятия

Под временным рядом понимается последовательность наблюдений значений некоторой переменной, произведенных через равные промежутки времени. Если принять длину такого промежутка за единицу времени (год, квартал, день и т.п), то можно считать, что последовательные наблюдения $x_1, ..., x_n$ произведены в моменты $t = 1, ..., n$.

Основная отличительная особенность статистического анализа временных рядов состоит в том, что последовательность наблюдений $x_1, ..., x_n$ рассматривается как реализация последовательности, вообще говоря, *статистически зависимых* случайных величин $X_1, ..., X_n$, имеющих некоторое совместное распределение с функцией распределения

$$F(v_1, v_2, ..., v_n) = P\{X_1 < v_1, X_2 < v_2, ..., X_n < v_n\}$$

Мы будем рассматривать в основном временные ряды, у которых совместное распределение случайных величин $X_1, ..., X_n$ имеет совместную плотность распределения $p(x_1, x_2, ..., x_n)$.

Чтобы сделать задачу статистического анализа временных рядов доступной для практического решения, приходится так или иначе ограничивать класс рассматриваемых моделей временных рядов, вводя те или иные предположения относительно структуры ряда и структуры его вероятностных характеристик. Одно из таких ограничении предполагает *стационарность* временного ряда.

Ряд x_t, $t = 1, ...n$ называется **строго стационарным** (или **стационарным в узком смысле**), если для любого m $(m < n)$ совместное распределение вероятностей случайных величин $X_{t_1}, ..., X_{t_m}$, такое же как и для $X_{t_1+\tau}, ..., X_{t_m+\tau}$, при любых $t_1, ..., t_m$ и τ, таких, что $1 \leq t_1, ..., t_m \leq n$ и $1 \leq t_1 + \tau, ..., t_m + \tau \leq n$, т.е функция плотности при сдвиге во времени не изменится

$$P\{x_{t_1}, ..., x_{t_m}\} = P\{x_{t_1+\tau}, ..., x_{t_m+\tau}\}$$

В частности, при $m = 1$ из предположения о строгой стационарности временного ряда x_t следует, что закон распределения вероятностей случайной величины X_t не зависит от t, а значит, не зависят от t и все его основные числовые характеристики (если, конечно, они существуют), в том числе: математическое ожидание $E(X_t) = \mu$ и дисперсия $D(X_t) = \sigma^2$.

Значение μ определяет постоянный уровень, относительно которого колеблется анализируемый временной ряд x_t, а постоянная σ характеризует размах этих колебаний.

Как мы уже говорили, одно из главных отличий последовательности наблюдений, образующих временной ряд, заключается в том, что члены временного ряда являются, вообще говоря, статистически взаимозависимыми. Степень тесноты статистической связи между случайными величинами X_t и $X_{t+\tau}$ может быть измерена парным коэффициентом корреляции

$$Corr(X_t, X_{t+\tau}) = \frac{Cov(X_t, X_{t+\tau})}{\sqrt{D(X_t)}\sqrt{D(X_{t+\tau})}}$$

где

$$Cov(X_t, X_{t+\tau}) = E[(X_t - E(X_t)(X_{t+\tau} - E(X_{t+\tau}))].$$

Если ряд стационарный, то значение $Cov(X_t, X_{t+\tau})$ не зависит от t и является функцией только от τ; мы будем использовать для него обозначение $\gamma(\tau)$:

$$\gamma(\tau) = Cov(X_t, X_{t+\tau}).$$

В частности,

$$D(X_t) = Cov(X_t, X_t) \equiv \gamma(0).$$

Соответственно, для стационарного ряда и значение коэффициента корреляции $Corr(X_t, X_{t+\tau})$ зависит только от τ; мы будем использовать для него обозначение $\rho(\tau)$, так что

$$\rho(\tau) = Corr(X_t, X_{t+\tau}) = \gamma(\tau)/\gamma(0).$$

В частности, $\rho(0) = 1$.

Практическая проверка строгой стационарности ряда x_t на основании наблюдения значений $x_1, x_2, ..., x_n$ в общем случае затруднительна.

В связи с этим под стационарным рядом на практике часто подразумевают временной ряд x_t, у которого

- $E(X_t) \equiv \mu$
- $D(X_t) \equiv \sigma^2$
- $Cov(X_t, X_{t+\tau} = \gamma(\tau)$ для любых t и τ.

Ряд, для которого выполнены указанные три условия, называют **стационарным в широком смысле (слабо стационарным, стационарным второго порядка** или **ковариационно стационарным)**.

Если ряд является стационарным в широком смысле, то он не обязательно является строго стационарным. В то же время, и строго стационарный ряд может не быть стационарным в широком смысле просто потому, что у него могут не существовать математическое ожидание и/или дисперсия. (В отношении последнего примером может служить случайная выборка из распределения Коши.) Кроме того, возможны ситуации, когда указанные три условия выполняются, но, например $E(X_t^3)$ зависит от t.

Ряд x_t $t = 1, ..., n$ называется **гауссовским**, если совместное распределение случайных величин $X_1, ..., X_n$ является n-мерным нормальным распределением. $X_t \sim N(\mu, \sigma^2)$. Для гауссовского ряда понятия стационарности в узком и широком смысле совпадают.

В дальнейшем, говоря о стационарности некоторого ряда x_t, мы (если не оговаривается противное) будем иметь в виду, что этот ряд стационарен в широком смысле (так что у него существуют математическое ожидание и дисперсия).

Итак, пусть x_t - стационарный ряд с $E(X_t) \equiv \mu, D(X_t) \equiv \sigma^2$ и $\rho(\tau) = Corr(X_t, X_{t+\tau})$. Поскольку в данном случае коэффициент $\rho(\tau)$ измеряет корреляцию между членами одного и того же временного ряда, его принято называть **коэффициентом автокорреляции** (или просто **автокорреляцией**). По той же причине о ковариациях $\gamma(\tau) = Cov(X_t, X_{t+\tau})$ говорят как об **автоковариациях**. При анализе изменения величины $\rho(\tau)$ в зависимости от значения τ принято говорить об **автокорреляционной функции** $\rho(\tau)$. Автокорреляционная функция безразмерна, т.е. не зависит от масштаба измерения анализируемого временного ряда. Ее значения могут изменяться в пределах от -1 до +1; при этом $\rho(0) = 1$. Кроме того, из стационарности ряда x_t следует, что $\rho(-\tau) = \rho(\tau)$, так что при анализе поведения автокорреляционных функций обычно ограничиваются рассмотрением только неотрицательных значений τ.

График зависимости $\rho(\tau)$ от τ часто называют **коррелограммой**. Он может использоваться для характеризации некоторых свойств механизма, порождающего временной ряд. Для дальнейшего заметим, что если x_t - стационарный временной ряд и c - некоторая постоянная, то временные ряды x_t и $(x_t + c)$ имеют одинаковые коррелограммы.

Если предположить, что временной ряд описывается моделью стационарного гауссовского процесса, то полное описание совместного распределения случайных величин $X_1, ..., X_n$ требует задания $n + 1$ параметров: $\mu, \gamma(0), \gamma(1), ..., \gamma(n-1)$ или $(\mu, \gamma(0), \rho(1), ..., \rho(n-1))$. Это намного меньше, чем без требования стационарности, но все же больше чем, количество наблюдений. В связи с этим, даже для стационарных гауссовских временных рядов приходится производить дальнейшее упрощение модели с тем, чтобы ограничить количество параметров, подлежающих оцениванию по имеющимся наблюдениям. Мы переходим теперь к рассмотрению некоторых простых по структуре временных рядов, которые, в то же время, полезны для описания эволюции во времени многих реальных экономических показателей.

2 Процесс белого шума

Процессом беого шума ("белым шумом "чисто случайным временным рядом") называют стационарный временной ряд x_t, для которого

$$E(X_t) \equiv 0, D(X_t) \equiv \sigma^2 > 0$$

и

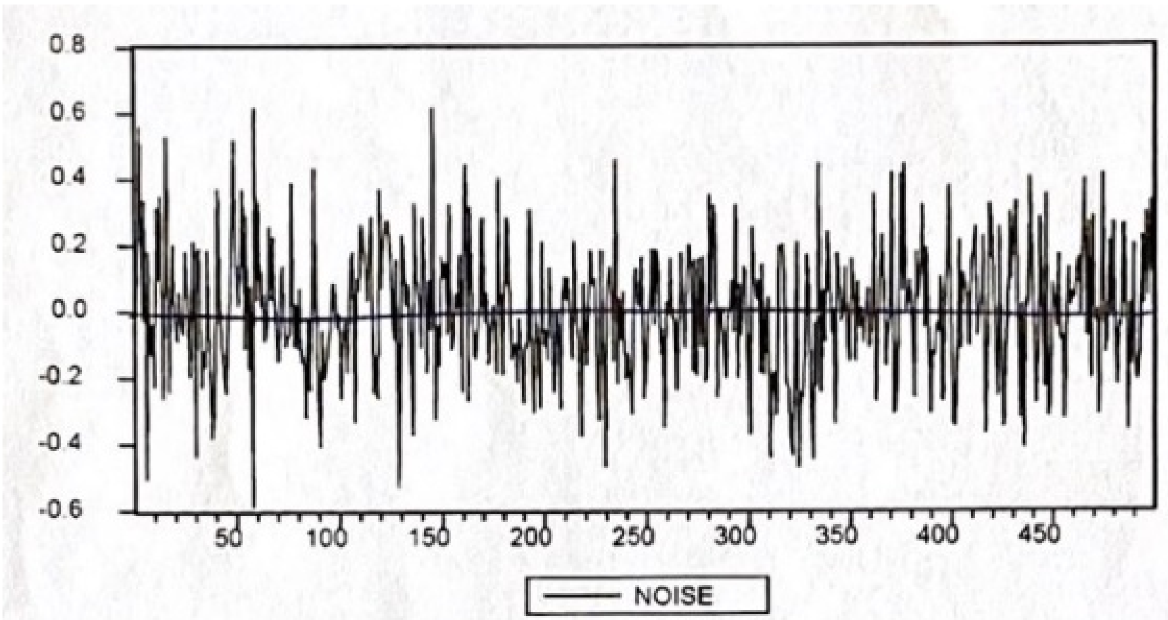
$$\rho(\tau) = 0, \text{ при } \tau \neq 0.$$

Последнее означает, что при $t \neq s$ случайные величины X_t и X_s, соответствующие наблюдениям процесса белого шума в моменты t и s, некоррелированы.

В случае, когда X_t имеет нормальное распределение, случайное величины $X_1, ...X_n$ взаимно независимы и имеют одинаковое нормальное распределение $N(0, \sigma^2)$, образуя случайную выборку из этого распределения, т.е. $X_t \sim i.i.d.N(0, \sigma^2)$. Такой ряд называют **гауссовским белым шумом**.

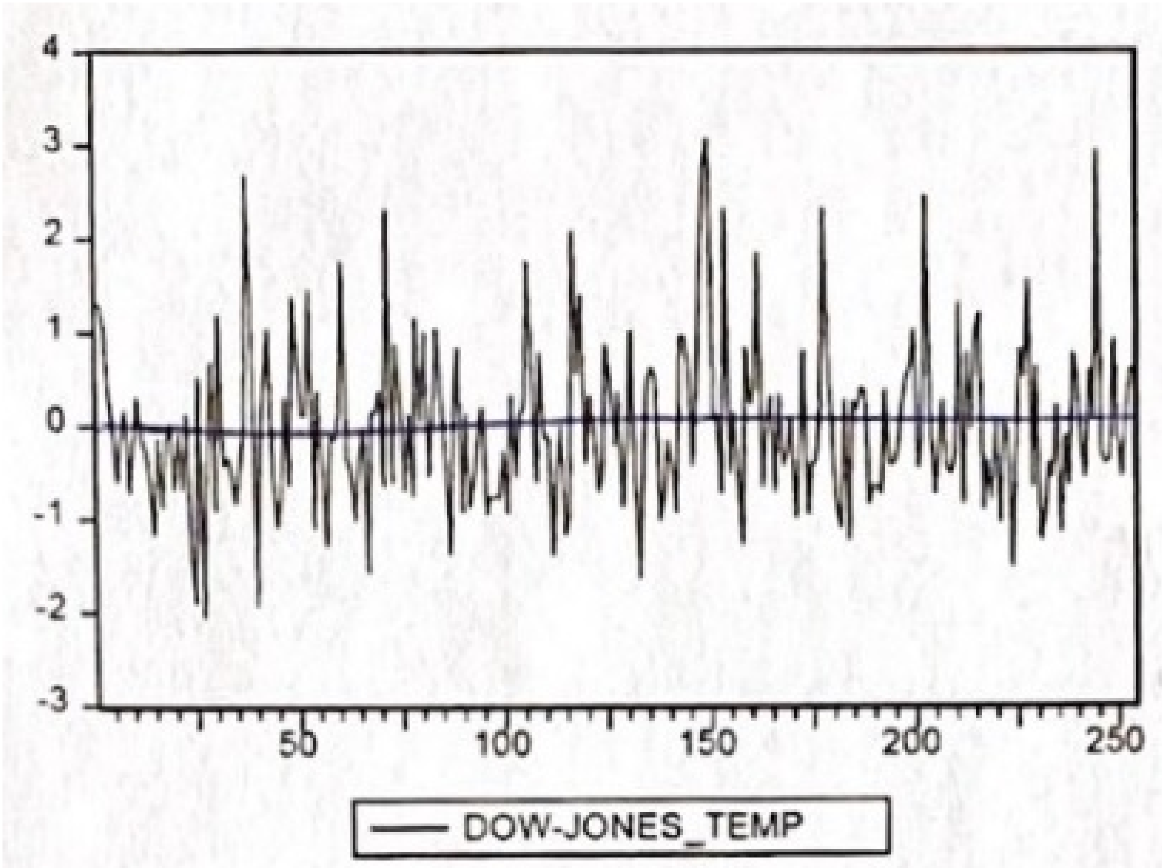
В то же время, в общем случае, даже если некоторые случайные величины $X_1, ..., X_n$ взаимно независимы и имеют одинаковое распределение, то это еще не означает, что они образуют процесс белого шума, т.к. случайная величина X_t может просто не иметь математического ожидания и/или дисперсии (в качестве примера мы опять можем указать на распределение Коши).

Временной ряд, соответствующий процессу белого шума, ведет себя крайне нерегулярным образом из-за некоррелированности при $t \neq s$ случайны величин X_t и X_s. Это иллюстрирует приводимый ниже график смоделированной реализации гауссовского процесса белого шума (NOISE) с $D(X_t) \equiv 0.04$.



В связи с этим процесс белого шума не годится для непосредственного моделирования эволюции большинства временных рядов, встречающихся в экономике. В то же время, как мы увидим ниже, такой процесс является базой для построения более реалистичных моделей временных рядов, порождающих "более гладкие"траектории ряда. В связи с частым использованием процесса белого шума в дальнейшем изложении, мы будем отличать этот процесс от других моделей временных рядов, используя для него обозначение ε_t.

В качестве примера ряда, траектория которого похожа на реализацию процесса белого шума, можно указать, например, на ряд, образованный значениями темпов изменения (прироста) индекса Доу - Джонса в течение 1984 года (дневные данные). График этого ряда имеет вид



Заметим, однако, что здесь наблюдается некоторая асимметрия распределения вероятностей значений x_t (скошенность этого распределения в сторону положительных значений), что исключает описание модели этого ряда как гауссовского белого шума.

3 Процесс авторегрессии

Одной из широко используемых моделей временных рядов является **процесс авторегрессии (модель авторегрессии)**. В своей простейшей форме модель авторегрессии описывает механизм порождения ряда следующим образом:

$$X_t = aX_{t-1} + \varepsilon_t, t = 1, ..., n,$$

где ε_t - процесс белого шума, для которого $E(\varepsilon_t) = 0$ и $D(\varepsilon_t) = \sigma^2 > 0$, X_0 - некоторая случайная величина, а $a \neq 0$ - некоторый постоянный коэффициент.

При этом

$$E(X_t) = aE(X_{t-1}),$$

так что рассматриваемый процесс может быть стационарным только если $E(X_t) = 0$ для всех $t = 0, 1, ..., n$.

Далее,

$$X_t = aX_{t-1} + \varepsilon_t = a(aX_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = a^2 X_{t-2} + a\varepsilon_{t-1} + \varepsilon_t = ...$$

$$= a^t X_0 + a^{t-1}\varepsilon_1 + a^{t-2}\varepsilon_2 + ... + \varepsilon_t,$$

$$X_{t-1} = aX_{t-2} + \varepsilon_{t-1} = a^{t-1}X_0 + a^{t-2}\varepsilon_1 + a^{t-3}\varepsilon_2 + ... + \varepsilon_{t-1},$$
$$X_{t-2} = aX_{t-3} + \varepsilon_{t-2} = a^{t-2}X_0 + a^{t-3}\varepsilon_1 + a^{t-4}\varepsilon_2 + ... + \varepsilon_{t-2}$$

$$...$$

$$X_1 = aX_0 + \varepsilon_1.$$

Если случайная величина X_0 не коррелирована со случайными величинами $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$, то отсюда следует, что

$$Cov(X_0, \varepsilon_1) = 0, Cov(X_1, \varepsilon_2) = 0, ..., Cov(X_{t-1}, \varepsilon_t) = 0$$

и

$$D(X_t) = D(aX_{t-1} + \varepsilon_t) = a^2 D(X_{t-1}) + D(\varepsilon_t), t = 1, ..., n.$$

Предполагая, наконец, что
$D(X_0) = D(X_t) = \sigma_X^2$ для всех $t = 1, ..., n$,

находим:

$$\sigma_X^2 = a^2\sigma_X^2 + \sigma_\varepsilon^2.$$

Последнее может выполняться только при выполнии условия $a^2 < 1$, т,е. $a < 1$.

При этом получаем выражение для σ_X^2

$$\sigma_X^2 = \frac{\sigma_\varepsilon^2}{1 - a^2}$$

Что касается автокорреляций и автоковариаций, то

$$Corr(X_t, X_{t+\tau}) = a^\tau$$

и

$$\begin{gathered}Cov(X_t, X_{t+\tau} = Cov(a^t X_0 + a^{t-1}\varepsilon_1 + a^{t-2}\varepsilon_2 + ... + \varepsilon_t,\\ a^{t+\tau}X_0 + a^{t+\tau-1}\varepsilon_1 + a^{t+\tau-2}\varepsilon_2 + ... + \varepsilon_{t+\tau}) =\\ = a^{2t+\tau}D(X_0) + a^\tau(1 + a^2 + ... + a^{2(t-1)})\sigma_\varepsilon^2 =\\ = a^\tau\left[\frac{a^{2t}\sigma_\varepsilon^2}{(1-a^2)} + \frac{(1-a^{2t})\sigma_\varepsilon^2}{(1-a^2)}\right] = \left[\frac{a^\tau}{(1-a^2)}\right]\sigma_\varepsilon^2.\end{gathered}$$

т.е при сделанных предположениях автоковариации и автокорреляции зависят только от того насколько разнесены по времени соответствующие наблюдения.

Таким образом, механизм порождения последовательных наблюдений, заданный соотношениями
$X_t = aX_{t-1} + \varepsilon_t, t = 1, ..., n,$
порождает стационарный временной ряд, если

- $a < 1$;
- случайная величина X_0 не коррелирована со случайными величинами $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$;
- $E(X_0) = 0$;
- $D(X_0) = \sigma_\varepsilon^2/(1-a^2)$.

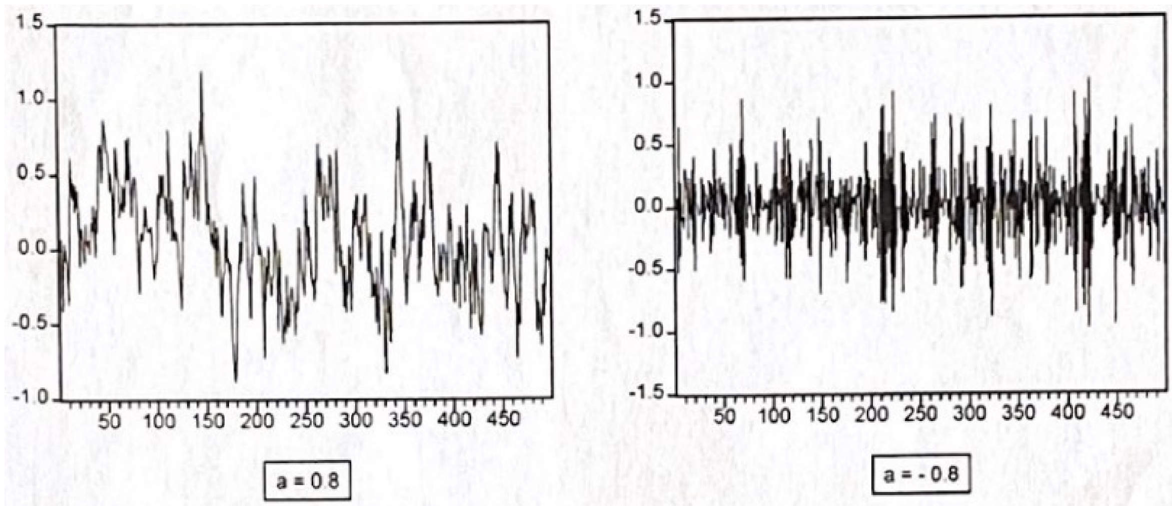
При этом

- $E(X_t) = 0$;
- $Corr(X_t, X_{t+\tau} = \rho(\tau) = a^\tau$;
- $Cov(X_t, X_{t+\tau} = \left[\dfrac{a^\tau}{(1-a^2)}\right]\sigma_\varepsilon^2$;
- $D(X_t) = \sigma_x^2 = \dfrac{\sigma_\varepsilon^2}{1-a^2}$

Рассмотренная модель порождает (при указанных условиях) стационарный ряд, имеющий нулевое математическое ожидание. Однако ее можно легко распространить и на временные ряды y_t с ненулевым математическим ожиданием $E(Y_t) = \mu$, полагая, что указанная модель относится к центрированному ряду $X_t = Y_t - \mu$:
$Y_t - \mu = a(Y_{t-1} - \mu) + \varepsilon_t, t = 1, ..., n,$
так что
$Y_t = aY_{t-1} + \sigma + \varepsilon_t, t = 1, ..., n,$
где
$\sigma = \mu(1 - a).$

Поэтому без ограничения общности можно обойтись в текущем рассмотрении моделями авторегрессии, порождающими стационарный процесс с нулевым средним.

Продолжая рассмотрение для ранее определенного процесса X_t (с нулевым математическим ожиданием), заметим, что для него
$\gamma(1) = E(X_t \cdot X_{t-1}) = E[(aX_{t-1} + \varepsilon_t) \cdot X_{t-1}] = a\gamma(0),$
так что
$\rho(1) = \gamma(1)/\gamma(0) = a,$
и при значениях $a > 0$, близких к 1, между соседними наблюдениями имеется сильная положительная корреляция, что обеспечивает более гладкий характер поведения траекторий ряда по сравнению с процессом белого шума. При $a < 0$ процесс авторегрессии, напротив, имеет менее гладкие реализации, поскольку в этом случае проявляется тенденция чередования знаков последовательных наблюдений. Следующие два графика демонстрируют поведение смоделированных реализаций временных рядов, порожденных моделями авторегрессии $X_t = aX_{t-1} + \varepsilon_t$ с $\sigma_\varepsilon^2 = 0.2$ при $a = 0.8$ (первый график) и $a = -0.8$ (второй график).



Теперь мы должны обратить внимание на следующее важное обстоятельство. В практических ситуациях "стартовое"значение $X_0 = x_0$, на основе которого в соответствии с соотношением $X_t = aX_{t-1} + \varepsilon_t$ строятся последующие значения ряда X_t, может относиться к концу предыдущего периода, на котором просто в силу других эконометрических условий эволюция соответствующего экономического показателя следует иной модели, например, модели $X_t = aX_{t-1} + \varepsilon_t$ с другими значениями a и σ_ε^2. Более того, статистические данные о поведении ряда до момента $t = 0$ могут отсутствовать вовсе, так что значение x_0 является просто некоторой наблюдаемой числовой величиной. В обоих случаях ряд X_t уже не будет стационарным даже при $a < 1$. Рассмотрим подробнее характеристики и поведение ряда в таких ситуациях.

Если не конкретизировать модель, в соответствии с которой порождались наблюдения до момента $t = 1$, то значение x_0 можно рассматривать как фиксированное. При этом
$X_t = a^t x_0 + a^{t-1}\varepsilon_1 + a^{t-2}\varepsilon_2 + ... + \varepsilon_t$,
$E(X_t) = a^t x_0 + a^{t-1}E(\varepsilon_1) + a^{t-2}E(\varepsilon_2) + ... + E(\varepsilon_t) = a^t x_0$,
$D(X_t) = (a^{2(t-1)} + a^{2(t-2)} + ... + 1)\sigma_\varepsilon^2 = \left[\dfrac{(1-a^{2t})}{(1-a^2)}\right]\sigma_\varepsilon^2$,
$Cov(X_t, X_{t+\tau}) = Cov(X_t - a^t x_0, X_{t+\tau} - a^{t+\tau}x_0) = a^\tau(1+a^2+...+a^{2(t-1)})\sigma_\varepsilon^2$

$$= a^\tau \left[\frac{(1-a^{2t})}{(1-a^2)}\right]\sigma_\varepsilon^2,$$

так что и математическое ожидание и дисперсия случайной величины X_t, а также ковариация $Cov(X_t, X_{t+\tau})$ зависят от t.

В то же время, если $a < 1$, то при $t \longrightarrow \infty$ получаем $E(X_t) \longrightarrow 0, D(X_t) \longrightarrow \sigma_\varepsilon^2/(1-a^2), Cov(X_t, X_{t+\tau}) \longrightarrow a^\tau[\sigma_\varepsilon^2/(1-a^2)]$, т.е. при $t \longrightarrow \infty$ значения математического ожидания и дисперсии случайной величины X_t, а также автоковариации $Cov(X_t, X_{t+\tau})$ стабилизируются, приближаясь к своим предельным значениям.

С этой точки зрения, условие $a < 1$ можно трактовать как условие стабильности ряда, порождаемого моделью $X_t = aX_{t-1} + \varepsilon_t$ при фиксированном значении $X_0 = x_0$. Рассмотрим в этой ситуации наряду с только что исследованным рядом X_t,

$$X_t = a^t x_0 + \sum_{k=0}^{t-1} a^k \varepsilon_{t-k},\ a < 1,$$

ряд порождаемый моделью

$$\widetilde{X}_t = \sum_{k=0}^{\infty} a^k \varepsilon_{t-k}.$$

Имеем:

$$\widetilde{X}_t - X_t = -a^t x_0 + \sum_{k=0}^{\infty} a^k \varepsilon_{t-k};$$

при $t \longrightarrow \infty$

$$a^t x_0 \longrightarrow 0 \text{ и } E\left|\sum_{k=0}^{\infty} a^k \varepsilon_{t-k}\right|^2 = \sigma_\varepsilon^2 \sum_{k=t}^{\infty} a^{2k} \longrightarrow 0.$$

$$\lim_{t \longrightarrow \infty} X_t = \widetilde{X}_t$$

Таким образом, ряд $\widetilde{X}_t$ является предельным для X_t; ряд X_t "выходит на режим"$\widetilde{X}_t$ при $t \longrightarrow \infty$. При этом выход ряда X_t на режим $\widetilde{X}_t$ происходит быстрее, чем ближе X_0 и a к нулю.

Для ряда $\widetilde{X}_t$

$$E(\widetilde{X}_t) = E\left(\sum_{k=0}^{\infty} a^k \varepsilon_{t-k}\right) = \sum_{k=0}^{\infty} a^k E(\varepsilon_{t-k}) = 0,$$

$$D(\widetilde{X}_t) = D\left(\sum_{k=0}^{\infty} a^k \varepsilon_{t-k}\right) = \sum_{k=0}^{\infty} a^k D(\varepsilon_{t-k}) = \sigma_\varepsilon^2 \sum_{k=0}^{\infty} a^{2k} = \frac{\sigma_\varepsilon^2}{1-a^2},$$

$$Cov(\widetilde{X}_t, \widetilde{X}_{t+\tau}) = E\left[\left(\sum_{k=0}^{\infty} a^k \varepsilon_{t-k}\right), \left(\sum_{k=0}^{\infty} a^k \varepsilon_{t+\tau-k}\right)\right] = a^\tau \left(\sum_{k=0}^{\infty} a^{2k} E(\varepsilon_{t-k}^2)\right) =$$

$$= a^\tau \frac{\sigma_\varepsilon^2}{1-a^2}$$

так что $\widetilde{X}_t$ - стационарный ряд (в широком смысле). Кроме того,

$$\widetilde{X}_{t-1} = \frac{1}{a} \sum_{k=1}^{\infty} a^k \varepsilon_{t-k},$$

так что

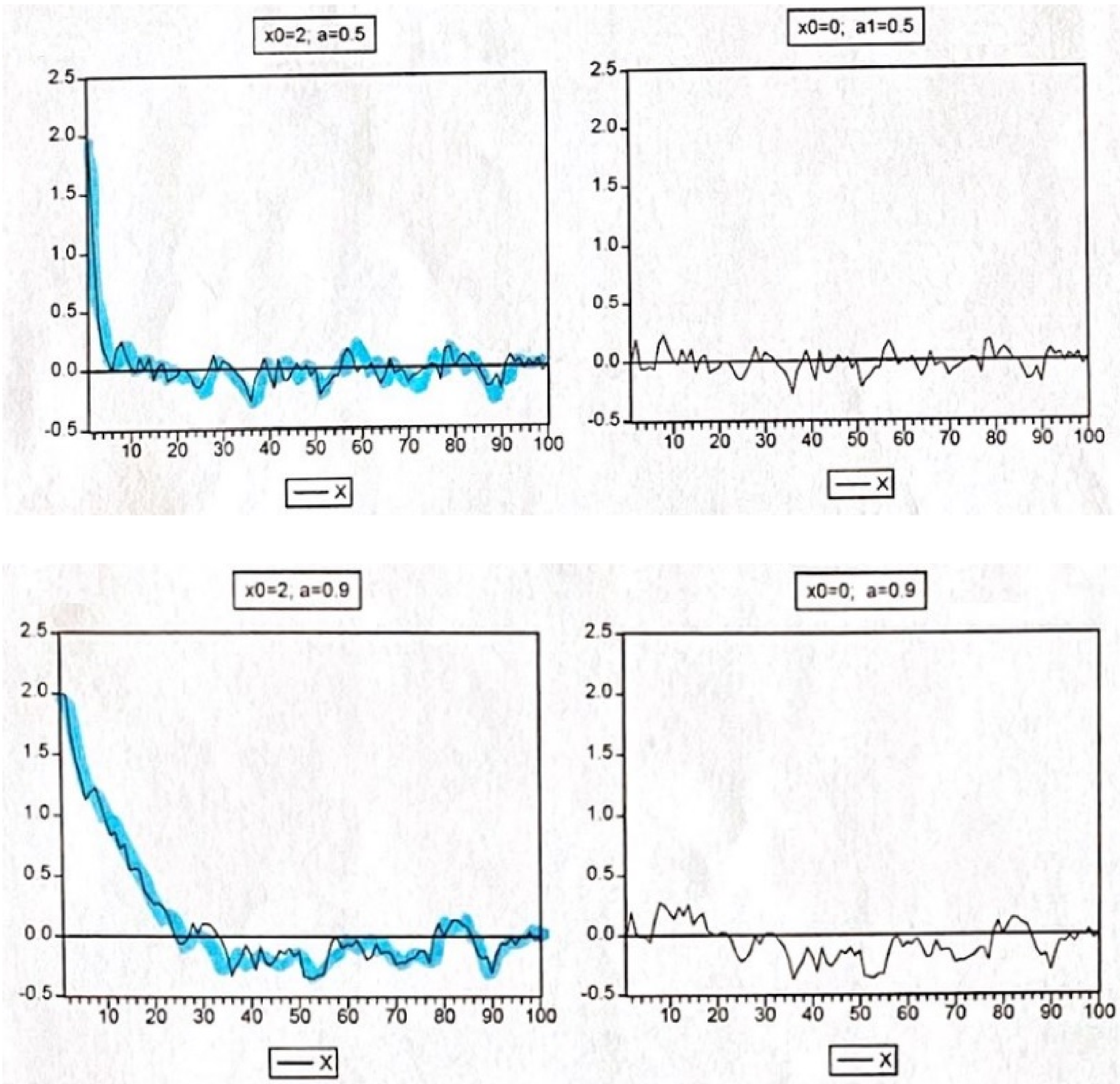
$$a\widetilde{X}_{t-1} + \varepsilon_t = \sum_{k=1}^{\infty} a^k \varepsilon_{t-k} = \widetilde{X}_t,$$

т.е. $\widetilde{X}_t$ удовлетворяет соотношению

$$\widetilde{X}_t = a\widetilde{X}_{t-1} + \varepsilon_t.$$

Поскольку ε_t не входит в правую часть выражений для $\widetilde{X}_{t-1}, \widetilde{X}_{t-2}, ...,$ то случайная величина ε_t не коррелирована с $\widetilde{X}_{t-1}, \widetilde{X}_{t-2}, ...,$ т.е. ε_t является инновацией (обновлением). В итоге получаем, что $\widetilde{X}_t$ - стационарный процесс авторегрессии певрого порядка, и фактически именно этот процесс имеется в виду, когда говорят о стационарном процессе AR(1).

Проиллюстрируем сказанное выше с помощью смоделированных реализацией ряда x_t, порожденных моделью $X_t = aX_{t-1} + \varepsilon_t$, с $\sigma_\varepsilon^2 = 0.2$ и различными значениями коэффициента a и стартового значения x_0.



Рассмотренную только что модель $X_t = aX_{t-1} + \varepsilon_t$ называют процессом авторегрессии первого порядка. **Процесс авторегрессии порядка p** (в кратком обозначении - **AR(p)**) определяется соотношениями

$X_t = a_1 X_{t-1} + a_2 X_{t-2} + ... + a_p X_{t-p} + \varepsilon_t, a_p \neq 0,$
где ε_t - процесс белого шума с $D(\varepsilon_t) = \sigma_\varepsilon^2$. Для простоты мы будем теперь сразу полагать, что $Cov(X_{t-s}, \varepsilon_t) > 0$ для всех $s > 0$; при этом говорят, что случайные величины ε_t образуют **инновационную (обновляющую) последовательность**, а случайная величина ε_t называется **инновацией** для наблюдения в момент t. Такая терминология

объясняется тем, что наблюдаемое значение ряда в момент t получается здесь как линейная комбинация p предшествующих значений этого ряда плюс не коррелированная с этими предшествующими значениями случайная составляющая ε_t, отражающая обновленную информацию, скажем, о состоянии экономики, на момент t, влияюущую на наблюдаемое значение X_t.

При рассмотрении процессов авторегрессии и некоторых других моделей удобно использовать **оператор запаздывания L (lag operator)**, который воздействует на временной ряд и определяется соотношением

$LX_t = X_{t-1}$
в некоторых руководствах его называют оператором обратного сдвига и используют для него обозначение **B (backshift operator)**.

Если оператор запаздывания применяется k раз, что обозначается как L^k, то это дает в результате

$$L^k X_t = X_{t-k}$$

Выражение

$$a_1 X_{t-1} + a_2 X_{t-2} + ... + a_p X_{t-p} =$$

можно записать теперь в виде

$$= (a_1 L + a_2 L^2 + ... + a_p L^p) X_t$$

а соотношение, определяющее процесс авторегрессии p-го порядка, в виде

$$a(L) X_t = \varepsilon_t$$

где

$$a(L) = 1 - (a_1 L + a_2 l^2 + ... + a_p L^p).$$

Для того, чтобы такой процесс был стационарным, все корни алгребраического уравнения

$$a(z) = 0$$

т.е. корни z: $|z| > 1$ (вещественные и комплексные) должны лежать вне единичного круга $|z| \leq 1$. (В частности, для процесса AR(1) имеем

$a(z) = 1 - az$, уравнение $a(z) = 0$ имеет корень $z = 1/a$, и условие стационарности $z > 1$ равносильно уже знакомому нам условию $a < 1$.) При этом **решение уравнения** $a(L)X_t = \varepsilon_t$, можно представить в виде

$$X_t = \frac{1}{a(L)}\varepsilon_t = \sum_{j=0}^{\inf} b_j \varepsilon_{t-j}$$

где $\sum_{j=0}^{\inf} |b_j| < \inf$ откуда, в частности, следует, что

$$E(X_t) = E\left(\sum_{j=0}^{\inf} b_j \varepsilon_{t-j}\right) = \sum_{j=0}^{\inf} b_j E(\varepsilon_{t-j}) = 0$$

Стационарный процесс $AR(p)$ с нулевым математическим ожиданием μ удовлетворяет соотношению

$$a(L)(X_t - \mu) = \varepsilon_t$$

или

$$a(L)X_t = \delta + \varepsilon_t$$

где

$$\delta = a(L)\mu = \mu(1 - a_1 - a_2 - ... - a_p) = \mu a(1)$$

При этом **решение уравнения** $a(L)(X_t - \mu) = \varepsilon_t$ имеет вид

$$X_t = \mu + \frac{1}{a(L)}\varepsilon_t$$

Таким образом, если стационарный процесс $AR(p)$ задан в виде $a(L)X_t = \delta + \varepsilon_t$, то следует помнить о том, что в этом случае математическое ожидание этого процесса равно не δ, а

$$E(x_t) = \mu = \frac{\delta}{(1 - a_1 - a_2 - ... - a_p)}$$

Для процесса $AR(1)$ имеем $a(L) = 1 - aL$, так что (**вне зависимости от того, равно μ нулю или нет**).

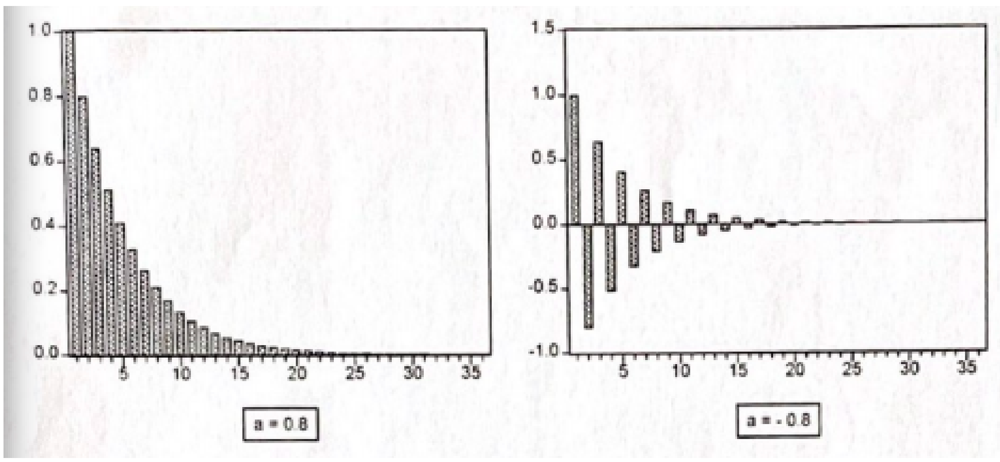
$$X_t - \mu = (\frac{1}{aL}\varepsilon_t) = (1 + aL + a^2L^2 + ...)\varepsilon_t = \varepsilon_t + a\varepsilon_{t-1} + a^2\varepsilon_{t-2} + ...$$

Из последнего выражения сразу видно, что

$$\rho(k) = Corr(X_t, X_{t+k}) = a^k \text{ , } k = 0, 1, 2, ...$$

При $0 < a < 1$ коррелограмма (график функции $\rho(k)$ для $k = 0, 1, 2, ...$) отражает **показательное убывание** корреляции с возрастанием интервала между наблюдениями; при $-1 < a < 0$ (коррелограмма имеет **характер затухающей косинусоиды**).

Сравним поведение коррелограммы стационарного процесса $AR(1)$ при $a = \pm 0.8$



Коррелограмма процесса $AR(p)$ при $p > 1$ имеет более сложную форму, зависящую от расположения (на комплексной плоскости) корней уравнения $a(z) = 0$. Однако для больших значений k автокорреляции $\rho(k)$ хорошо апроксимируется значением $A\theta^k$, где $\theta = 1/z_{\min}$ и $z_{\min}$ - наименьший по абсолютной величине корень уравнения $a(z) = 0$, если этот корень являетя вещественным и положительным, или заключена в интервале $\pm A\theta^k$ в противном случае. Здесь $A > 0$ - некоторая постоянная, определяемая коэффицентами $a_1, ..., a_p$

Если умножить на $X_{t-k}(k > 0)$ обе части соотношения, определяющего процесс $AR(p)$, и после этого взять от обеих частей математическое ожидание, то получим соотношение

$$\gamma(k) = a_1\gamma(k-1) + a_2\gamma(k-2) + ... + a_p\gamma(k-p) \text{ , } k > 0$$

$$\gamma(k) = Cov(X_t, X_{t-k})$$

Разделив обе части последнего на $\gamma(0)$, приходим к **системе Юла-Уокера**

$$\rho(k) = a_1\rho(k-1) + a_2\rho(k-2) + ... + a_p\rho(k-p)\ \textbf{,}\ k > 0$$

Эта система позволяет последовательно находить значения автокорреляции и дает возможность, используя первые p уравнений, выразить коэффиценты a_j через значения первых p автокорреляций, что можно непосредственно использовать при подборе модели авторегрессии к реальным статистическим данным (см. разделы 3.1 и 3.2)

4 Пример

Рассмотрим процесс авторегрессии $AR(2)$

$$X_r = 4.375 + 0.25X_{t-1} - 0.125X_{t-2} + \varepsilon_t$$

Уравнение $a(z) = 0$ принимает в этом случае вид

$$1 - 0.25z + 0.125z^2 = 0 \textbf{ или } z^2 - 2z + 8 = 0$$

и имеет корни $z_{1,2} = 1 \pm i\sqrt{7}$. Оба корня по абсолютной величине больше единицы, так что процесс стационарный. Математичесое ожидание этого процесса равно

$$\mu\delta/(1 - a_1 - a_2) = 4.375/(1 - 0.25 - 0.125) = 5$$

так что траектории этого процесса флуктурирует вокруг уровня 5.

Для построения коррелограммы воспользуемся уравнением Юла-Уокера. У нас $p = 2$, так что

$$\rho(k) = 0.25\rho(k-1) - 0.125\rho(k-2)\ \textbf{,}\ k > 0$$

По определению, $\rho(0) = 1$. Для $\rho(1)$ имеем

$$\rho(1) = 0.25\rho(0) - 0.125\rho(-1) = 0.25 - 0.125\rho(1)$$

откуда находим:

$$\rho(1) = 0.25/(1 + 0.125) = 2/9 = 0.222$$

Далее последовательно находим:

$$\rho(2) = 0.25\rho(1) - 0.125\rho(0) = 0.25 * 0.222 - 0.125 = -0.069$$

$$\rho(3) = -0.045,\ \rho(4) = -0.003,\ \rho(5) = 0.005 \textbf{ и т.д.}$$

Корреляции даже между соседними наблюдениями очени малы, и можно ожидать, что поведение траекторий такого ряда не очень существенно от поведения реализаций процесса белого шума. Теоретическая коррелограмма расматриваемого процесса и смоделированная реализация этого процесса приведены ниже.