

Предположим, что мы определили единицу измерения времени (один день, один месяц или один год). Последовательность $S = (S_n)_{n \geq 0}$ – рыночная стоимость акции, обменный курс двух валют или что-либо иное. В этом учебнике мы будем придерживаться вероятностного подхода и следовать общепринятой аксиоматике теории вероятностей. То есть S – стохастическая последовательность, которая определена на некотором вероятностном пространстве (Ω, F, P) , где

Ω – пространство элементарных случайных событий (случаев) $\omega \in \Omega$ (в рассматриваемом контексте состояний рынка);

F – σ -алгебра подмножеств Ω – совокупность событий, наблюдаемых на рынке;

P – семейство вероятностных мер, возможно, параметрическое, на F .

Для различных финансовых теорий время и динамика являются неотъемлемой частью рассуждений, поэтому целесообразно рассмотреть поток σ -подалгебр $F_0 \subseteq F_1 \subseteq \dots \subseteq F_n \subseteq \dots \subseteq F$. Смысл введения потока, называемого в литературе также фильтрацией, заключается в том, чтобы в любой момент времени n оперировать только теми случайными событиями, которые «доступны» наблюдателю до момента времени n включительно. Например, до момента времени n включительно наблюдателю могут быть доступны стоимости тех или иных активов, начиная с некоторого начального момента до момента времени n .

Таким образом, базовой вероятностной моделью является модель: $(\Omega, F, (F_n)_{n \geq 0}, P)$, называемая фильтрованным стохастическим экспериментом.

Если рассматривать F_n как информацию, доступную к моменту времени n , то естественно считать, что последовательность S – адаптирована, то есть для любого момента времени n S_n – F_n -измеримы. Интерпретация S_n как цены в момент времени n приводит к тому, что $S_n > 0$.

Существует два наиболее распространенных способа представления временного ряда S . Первый способ: $S_n = S_{n-1} \exp(h_n)$, откуда $S_n = S_0 \exp(H_n)$,

где $H_n = \sum_{i=1}^n h_i$ и $h_n = \ln \frac{S_n}{S_{n-1}}$ – логарифмический возврат (return). Другой способ

представления: $S_n = (1 + \rho_n) S_{n-1}$, откуда $S_n = S_0 \prod_{i=1}^n (1 + \rho_i)$ и $\rho_n = \frac{\Delta S_n}{S_{n-1}}$.

Рассмотрим последовательность $R_n = \sum_{i=1}^n \rho_i$ и обозначим через

$E(R)_n = \prod_{i=1}^n (1 + \rho_i)$, $E(R)_0 = 1$. Определяемая этим выражением

последовательность называется стохастической экспонентой, порождаемой R .
Имеем: $S_n = S_0 E(R)_n$. Справедливы соотношения:

$$\ln E(R)_n = H_n,$$

$$H_n = \sum_{i=1}^n \ln(1 + \rho_i),$$

$$h_i = \ln(1 + \rho_i),$$

$$R_n = \sum_{i=1}^n (\exp(h_i) - 1).$$

При малых значениях ρ_n $h_n \approx \rho_n$, при этом ошибка $|h_n - \rho_n| \leq \frac{1}{2} \rho_n^2$.

Таким образом, представление 1 и представление 2 в одинаковой степени могут быть использованы для описания последовательности S .

Далее мы будем использовать первое представление, поскольку второе представление накладывает ограничение на ρ . Из того, что $S > 0$, непосредственно следует, что $\rho > -1$. В то же время различного рода ограничения могут существенно отразиться на сложности статистических задач.

Линейный прогноз

Определение. Гильбертовым пространством вещественнозначных случайных величин с конечным вторым моментом $L^2(\Omega, F, P)$ называется линейное пространство случайных величин с $E\xi = 0, E\xi^2 < \infty$ и скалярным произведением: $(\xi, \eta) = E\xi\eta = Cov(\xi, \eta)$.

Рассмотрим совокупность линейно-независимых случайных величин $\eta_1, \eta_2, \dots, \eta_n$, обозначим через L_η линейную оболочку, натянутую на случайные величины $\eta_1, \eta_2, \dots, \eta_n$.

Наилучшим линейным прогнозом случайной величины ξ по совокупности случайных величин $\eta_1, \eta_2, \dots, \eta_n$ назовем: $\hat{\xi} = \arg \min_{\zeta \in L_\eta} E(\xi - \zeta)^2$.

Так как $\zeta \in L_\eta$, то $\zeta = \sum_{i=1}^n a_i \eta_i$, следовательно, наилучший линейный прогноз: $\hat{\xi} = \arg \min_a E\left(\xi - \sum_{i=1}^n a_i \eta_i\right)^2$.

Минимум достигается тогда, когда разность $\xi - \sum_{i=1}^n a_i \eta_i$ ортогональна подпространству L_η . Таким образом, решение задачи сводится к решению системы линейных алгебраических уравнений $E\left(\xi - \sum_{i=1}^n a_i \eta_i\right) \eta_j = 0$. В

матричных обозначениях система уравнений представляется в виде: $CX = Y$,

$$X = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}, C_{i,j} = \text{Cov}(\eta_i, \eta_j), Y = \begin{pmatrix} \text{Cov}(\xi, \eta_1) \\ \text{Cov}(\xi, \eta_2) \\ \dots \\ \text{Cov}(\xi, \eta_n) \end{pmatrix}.$$

Дисперсия ошибки

$$D\Delta_\eta = E\left(\xi - \sum_{i=1}^n a_i \eta_i\right)^2 = D\xi - 2\sum_{i=1}^n a_i \text{Cov}(\xi, \eta_i) + \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\eta_i, \eta_j) a_i a_j.$$

Используем предыдущие обозначения, в результате получим $D\Delta_\eta = D\xi - 2(X, Y) + (CX, X)$. Откуда $D\Delta_\eta = D\xi - (C^{-1}Y, Y)$.

В последних выражениях (\dots) – скалярное произведение евклидова пространства.

Наилучшим прогнозом случайной величины ξ по совокупности случайных величин $\eta_1, \eta_2, \dots, \eta_n$ назовём: $\hat{\xi} = \arg \min_{\zeta \in F_\eta} E(\xi - \zeta)^2$, где $F_\eta = \sigma(\eta_1, \dots, \eta_n)$ – минимальная сигма-подалгебра F , содержащая события

$$A_b = \{\omega \in \Omega : \eta(\omega) \leq b\}; \quad \eta = \begin{pmatrix} \eta_1(\omega) \\ \dots \\ \eta_n(\omega) \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}; \quad \zeta \in F_\eta \text{ означает, что } \zeta - F_\eta -$$

измерима. Из телескопического свойства условного математического ожидания $E(\xi - \zeta)^2 = E\left(E\left((\xi - \zeta)^2 / F_\eta\right)\right)$. Из неравенства

$$E\left(E\left((\xi - \zeta)^2 / F_\eta\right)\right) \leq E\left(E\left((\xi - E(\xi / F_\eta))^2 / F_\eta\right)\right), \forall \zeta \in F_\eta, \text{ следует неравенство}$$

$$E(\xi - \zeta)^2 \leq E\left(\xi - E(\xi / F_\eta)\right)^2, \forall \zeta \in F_\eta. \text{ Отсюда оптимальный прогноз: } \hat{\xi} = E(\xi / F_\eta).$$

Если совместный закон распределения $Low(\xi, \eta_1, \dots, \eta_n)$ – многомерный нормальный закон распределения, то по теореме о нормальной корреляции наилучший прогноз совпадает с наилучшим линейным прогнозом.

Гауссовы модели

Остановимся на проблеме описания распределения вероятностей последовательности S , что сводится к описанию распределения вероятностей последовательности H или последовательности h .

С точки зрения классической теории вероятностей и далеко продвинутой статистики нормального распределения было бы весьма привлекательным использование нормального закона распределения для описания поведения последовательности h . Поскольку для задания распределения в этом случае достаточно задания последовательности математических ожиданий $\mu_n = Eh_n$ и

ковариационной матрицы $Cov(h_n, h_m) = E h_n h_m - \mu_n \mu_m$. Принятие такого предположения существенно упрощает решение задачи прогнозирования h_{n+1} по h_1, h_2, \dots, h_n . Как отмечалось выше, наилучший прогноз:

$$\hat{h}_{n+1} = E(h_{n+1} | h_1, h_2, \dots, h_n) = \mu_{n+1} + \sum_{i=1}^n a_i (h_i - \mu_i).$$

Коэффициенты a_i удовлетворяют системе линейных алгебраических уравнений.

Ошибка прогноза $\Delta_{n+1} = h_{n+1} - \hat{h}_{n+1}$ распределена по нормальному закону распределения с математическим ожиданием $E\Delta = 0$ и дисперсией: $D\Delta_{n+1} = Dh_{n+1} - (C^{-1}Y, Y)$.

Теперь мы можем говорить о доверительном интервале для h_{n+1} при заданной доверительной вероятности $p_{дог}$. Величина h_{n+1} принадлежит доверительному интервалу: $[\hat{h}_{n+1} - k\sqrt{D\Delta_{n+1}}, \hat{h}_{n+1} + k\sqrt{D\Delta_{n+1}}]$,

где k находится из уравнения: $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^k \exp\left(-\frac{x^2}{2}\right) dx = \frac{1 + p_{дог}}{2}$.

Зная доверительный интервал для h_{n+1} , можно указать доверительный интервал для S_{n+1} : $[S_n \exp(\hat{h}_{n+1} - k\sqrt{D\Delta_{n+1}}), S_n \exp(\hat{h}_{n+1} + k\sqrt{D\Delta_{n+1}})]$.

Пусть C – диагональная матрица, тогда формулы существенно упрощаются. Коэффициенты $a_i = \frac{Cov(h_{n+1}, h_i)}{Dh_i}$ и дисперсия ошибки

$$D\Delta_{n+1} = Dh_{n+1} - \sum_{i=1}^n \frac{Cov^2(h_{n+1}, h_i)}{Dh_i}.$$

Равенство нулю ковариаций для нормального закона распределения равносильно независимости. В этом случае матрица C – диагональная матрица, а Y – нулевой вектор, и наилучший прогноз $\hat{h}_{n+1} = \mu_{n+1}$, дисперсия ошибки которого $D\Delta_{n+1} = Dh_{n+1}$. Если добавить еще более сильное требование одинаковой распределённости элементов последовательности h , то при таких предположениях цены S анализируются достаточно просто. В этом случае прогноз $\hat{h}_{n+1} = \mu$ и дисперсия ошибки $D\Delta_{n+1} = d$. То есть теряется возможность какого-либо прогноза будущих значений.

В реальности ситуация более благоприятна, поскольку многочисленные исследования временных финансовых последовательностей показывают, что последовательность h не является гауссовской последовательностью. Кроме этого, исследования показывают наличие зависимости между элементами последовательности h .

Более ёмким распределением является смесь гауссовских распределений. В простейшем случае плотность смеси:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \left(p_1 \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) + (1-p_1) \exp\left(-\frac{(x-b)^2}{2\sigma^2}\right) \right).$$

Это распределение интерпретируется формулой: $h_n = \delta_n b + (1-\delta_n)a + \sigma\varepsilon_n = (b-a)\delta_n + a + \sigma\varepsilon_n$, в которой δ – последовательность независимых и одинаково распределенных бинарных случайных величин, ε – последовательность независимых стандартных гауссовых величин, не зависящих от δ .

Следовательно, $H_n = (b-a)\xi_n + na + \sqrt{n}\sigma\varepsilon_n$, где случайная величина ξ_n распределена по закону Бернулли.

Данную модель можно рассматривать как модель Кокса-Росса-Рубинштейна с белым шумом.

Математическое ожидание и дисперсия h_n :
 $Eh_n = (b-a)p + a$, $Dh_n = (b-a)^2 p(1-p) + \sigma^2$.

Случайные величины в этой модели независимы и одинаково распределены, поэтому наилучший линейный прогноз: $\hat{h}_{n+1} = (b-a)p + a$, при этом дисперсия ошибки: $D\Delta_{n+1} = Dh_{n+1} = (b-a)^2 p(1-p) + \sigma^2$.

Отметим, что в этом случае наилучший линейный прогноз не совпадает с наилучшим прогнозом.

Данная модель допускает определенное развитие, в котором учитывается зависимость в последовательности h : $h_n = (b-a)\delta_n \delta_{n-1} + a\delta_{n-1} + \sigma\varepsilon_n$, при $\delta_0 = 1$.

Для H : $H_n = (b-a) \sum_{i=1}^n \delta_i \delta_{i-1} + a \sum_{i=1}^n \delta_i + \sqrt{n}\sigma\varepsilon_n$.

В этой модели $h_n = \begin{cases} \sigma\varepsilon_n, \text{ при } \delta_{n-1} = 0 \\ (b-a)\delta_n + a + \sigma\varepsilon_n, \text{ при } \delta_{n-1} = 1 \end{cases}$.

Назовём данную модель моделью Кокса-Росса-Рубинштейна с зависимыми значениями.

Математическое ожидание: $Eh_n = p((b-a)p + a)$.

Ковариация $Cov(h_{n+k}, h_n) = 0$ при $k > 1$. При $k = 1$:
 $Cov(h_{n+1}, h_n) = (b-a)p^2(1-p)((b-a)p + a)$. Дисперсия:

$Dh_{n+1} = p(1-p)(p(1+p)(b-a)^2 + 2(b-a)ap + a^2) + \sigma^2$.

Введем обозначения $Cv = Cov(h_{n+1}, h_n)$ и $Ds = Dh_{n+1}$. В этих обозначениях

матрица $C = \begin{pmatrix} Ds & Cv & 0 & \dots & \dots & 0 \\ Cv & Ds & Cv & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Cv & Ds \end{pmatrix}$, $Y = \begin{pmatrix} 0 \\ 0 \\ \dots \\ Cv \end{pmatrix}$. Представим

ковариационную матрицу C в виде произведения: $C = AA^T$,

где матрица $A = \begin{pmatrix} \alpha_{1,1} & 0 & \dots & 0 \\ \alpha_{2,1} & \alpha_{2,2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \alpha_{n-1,n} & \alpha_{n,n} \end{pmatrix}$. Элементы матрицы A удовлетворяют

уравнениям: $\alpha_{1,1} = \sqrt{Ds}$, $\alpha_{i+1,i} = \frac{Cv}{\alpha_{i,i}}$, $\alpha_{i+1,i+1} = \sqrt{Ds - \alpha_{i+1,i}^2}$, $i = 1, \dots, n-1$.

Запишем уравнение в виде: $AA^T X = Y$. Введем обозначение $Z = A^T X$. Имеем: $AZ = Y$, $A^T X = Z$.

Решение первого уравнения $Z = \begin{pmatrix} 0 \\ 0 \\ \dots \\ \frac{Cv}{\alpha_{n,n}} \end{pmatrix}$. Решение второго уравнения:

$$x_n = \frac{Cv}{\alpha_{n,n}^2}, x_{n-k} = -\frac{\alpha_{n-k+1, n-k} x_{n-k+1}}{\alpha_{n-k, n-k}}, k = 1, \dots, n-1.$$

Таким образом, наилучший линейный прогноз:

$$\hat{h}_{n+1} = p((b-a)p + a) + \sum_{i=1}^n x_i (h_i - p((b-a)p + a))$$

При этом дисперсия ошибки:

$$D\Delta_{n+1} = p(1-p) \left(p(1+p)(b-a)^2 + 2(b-a)ap + a^2 \right) + \sigma^2 - \frac{Cv^2}{\alpha_{n,n}^2}.$$

Потребуем, чтобы последовательность δ была марковской цепью, с не зависящими от n переходными вероятностями

$$P\left(\delta_n = 1 \middle/ \delta_{n-1} = 0\right) = r, P\left(\delta_n = 1 \middle/ \delta_{n-1} = 1\right) = s. \text{ Обозначим через } Q = \begin{pmatrix} 1-r & 1-s \\ r & s \end{pmatrix}$$

матрицу переходных вероятностей. Назовём данную модель моделью Кокса-Росса-Рубинштейна с марковской цепью.

Математическое ожидание $Eh_n = (b-a)P(\delta_n = 1) + a$. Вероятность

$$P(\delta_n = 1) = (Q^{n-1} P_1, I), \text{ где } P_1 = \begin{pmatrix} P(\delta_1 = 0) \\ P(\delta_1 = 1) \end{pmatrix}, I = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \text{ Собственные числа матрицы}$$

Q : $\lambda_1 = 1, \lambda_2 = s - r$. Соответствующие собственные векторы:

$$U_1 = \begin{pmatrix} 1-s \\ r \\ 1 \end{pmatrix}, U_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \text{ Таким образом, для матрицы } Q \text{ существует сингулярное}$$

разложение: $Q = U \Lambda U^{-1}$,

где $\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & s-r \end{pmatrix}, U = \begin{pmatrix} 1-s & -1 \\ r & 1 \end{pmatrix}$. Следовательно, $Q^k = U \Lambda^k U^{-1}$. Матрица

$$U^{-1} = \frac{r}{1+r-s} \begin{pmatrix} 1 & 1 \\ -1 & \frac{1-s}{r} \end{pmatrix}. \quad \text{Отсюда}$$

$$Eh_n = \frac{b-a}{1+r-s} \left(r + (s-r)^n (P(\delta_1=1)(1-s) - P(\delta_1=0)r) \right) + a.$$

Поскольку $|s-r| < 1$, то $\lim_{n \rightarrow \infty} (s-r)^n = 0$. Следовательно,

$$\lim_{n \rightarrow \infty} Eh_n = \frac{r(b-a)}{1+r-s} + a, \text{ причем скорость сходимости показательная.}$$

Вычислим ковариацию $Cov(h_n, h_m)$. Не нарушая общности, будем считать, что $m < n$. Прежде всего, определим математическое ожидание $Eh_n h_m = E((b-a)h_n + a)((b-a)h_m + a) =$

$$(b-a)^2 P\left(\frac{\delta_n=1}{\delta_m=1}\right) P(\delta_m=1) + (b-a)a(P(\delta_m=1) + P(\delta_n=1)) + a^2.$$

Вероятности $P\left(\frac{\delta_n=1}{\delta_m=1}\right) = (Q^{n-m} I, I), P(\delta_k=1) = (Q^k P_1, I)$. Отсюда

$$Eh_n h_m = (b-a)^2 (Q^{n-m} I, I) (Q^{m-1} P_1, I) + (b-a)a((Q^{m-1} P_1, I) + (Q^{n-1} P_1, I)) + a^2.$$

Далее будем считать, что последовательность $\delta = \dots, \delta_{-p}, \delta_{-p+1}, \dots, \delta_0, \delta_1, \dots, \delta_n, \dots$. Обозначим через $Q^\infty = \lim_k Q^k$. Для рассматриваемого варианта имеем:

$$Cov(h_n, h_m) = (b-a)^2 (Q^\infty P_1, I) ((Q^{n-m} I, I) - (Q^\infty I, I)). \text{ Тогда } Q^\infty = U \lim_{k \rightarrow \infty} \Lambda^k U^{-1}. \text{ Так}$$

$$\text{как } \lim_{k \rightarrow \infty} \Lambda^k = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ то } Q^\infty = \frac{r}{1+r-s} \begin{pmatrix} 1-s & 1-s \\ r & r \\ 1 & 1 \end{pmatrix}.$$

$$\text{Имеем: } Cov(h_n, h_m) = (b-a)^2 \frac{r(1-s)}{(1+r-s)^2} (s-r)^{n-m}.$$

Дисперсия: $Dh_n = \frac{r(1-s)}{1+r-s}$, что выглядит достаточно естественным, если

учесть, что в рассматриваемой ситуации

$P(\delta_n = 1) = \frac{r}{1+r-s}, P(\delta_n = 0) = 1 - \frac{r}{1+r-s} = \frac{1-s}{1+r-s}$. Отметим также, что

$$Eh_n = (b-a)P(\delta_n = 1) + a = \frac{r(b-a)}{1+r-s} + a.$$

Последовательности, обладающие свойством: $Cov(h_m, h_n) = f(|m-n|), Eh_n = const$, называются стационарными в широком смысле.

Соединение стационарности и гауссовости приводит к стационарности последовательности S в узком смысле.

Последовательность называется стационарной в узком смысле, если для любых n и k $Low(h_1, h_2, \dots, h_n) = Low(h_{k+1}, h_{k+2}, \dots, h_{k+n})$.

Далее мы будем считать, что последовательность h стационарна в широком смысле. Чтобы упростить выкладки, будем считать, что $Eh_n = 0$. Элементы последовательности принадлежат гильбертову пространству комплекснозначных случайных величин $H^2(\Omega, F, P)$ со скалярным произведением $(\xi, \zeta) = E\xi\bar{\zeta} = Cov(\xi, \zeta)$, $\bar{\zeta}$ – комплексно-сопряжённое число к ζ .

Обозначим через $R(k) = Cov(h_{n+k}, h_n) = Eh_{n+k}\bar{h}_n$, $\rho(k) = \frac{R(k)}{R(0)}$.

Естественно, что мы предполагаем, что $R(0) = Dh_n \neq 0$.

Функция $R(k)$ называется ковариационной функцией. Функция $\rho(k)$ называется корреляционной функцией. Из определения следует, что ковариационная функция является неотрицательно определенной, то есть для любых комплексных чисел a_1, a_2, \dots, a_m и любых $t_1, t_2, \dots, t_m \in \mathbf{Z}$, $m \geq 1$,

$$\sum_{i,j=1}^m R(t_i - t_j) a_i \bar{a}_j \geq 0.$$

В качестве упражнения доказать следующие свойства

ковариационной функции: $R(0) \geq 0, R(-k) = \bar{R}(k), |R(n)| \leq R(0), S |R(n) - R(m)| \leq 2R(0)(R(0) - \text{Re} R(n-m))$.

Пример 1. Последовательность $h_n = \xi \exp(i\lambda n)$, где $E\xi = 0$, $E|\xi|^2 < \infty$ и $\lambda \in [-\pi, \pi)$ является последовательностью, стационарной в широком смысле. Действительно, $Eh_n = 0, R(k) = E|\xi|^2 \exp(i\lambda k)$.

Пример 2. Последовательность $h_n = \sum_{j=1}^N z_j \exp(i\lambda_j n)$, где z_1, z_2, \dots, z_N –

ортогональные случайные величины с $Ez_j = 0$ и $E|z_j|^2 = \sigma_j^2 < \infty$, $-\pi \leq \lambda_1 < \lambda_2 < \dots < \lambda_N < \pi$, является стационарной. Действительно,

$$Eh_n = 0, R(k) = \sum_{j=1}^N \sigma_j^2 \exp(i\lambda_j k).$$

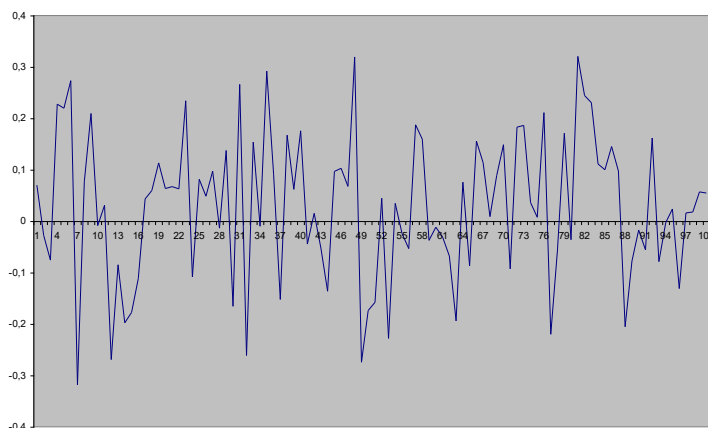


Рисунок 1. График компьютерной реализации последовательности h в случае модели Кокса-Росса-Рубинштейна с белым шумом при $a = -0.1, b = 0.1, P(\delta_n = 1) = 0.6, \sigma = 0.1$.

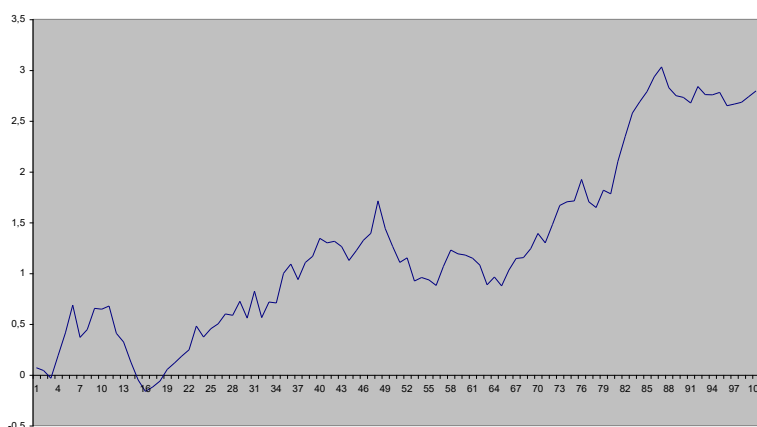


Рисунок 2. График компьютерной реализации последовательности H в случае модели Кокса-Росса-Рубинштейна с белым шумом при $a = -0.1, b = 0.1, P(\delta_n = 1) = 0.6, \sigma = 0.1$.

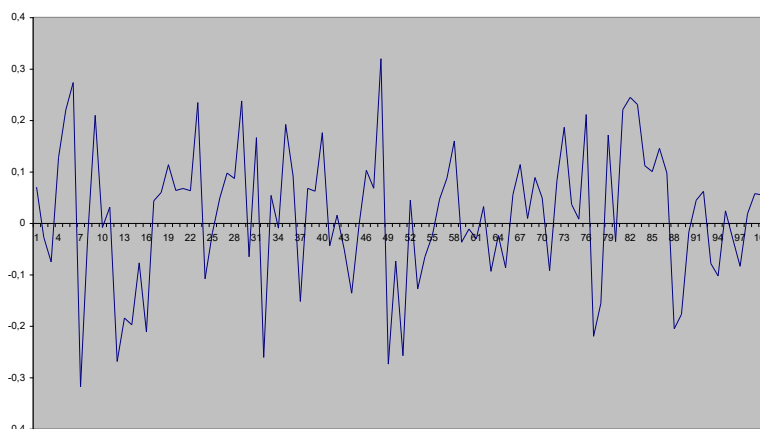


Рисунок 3. График компьютерной реализации последовательности h в случае модели Кокса-Росса-Рубинштейна с зависимыми значениями при $a = -0.1, b = 0.1, P(\delta_n = 1) = 0.6, \sigma = 0.1$.

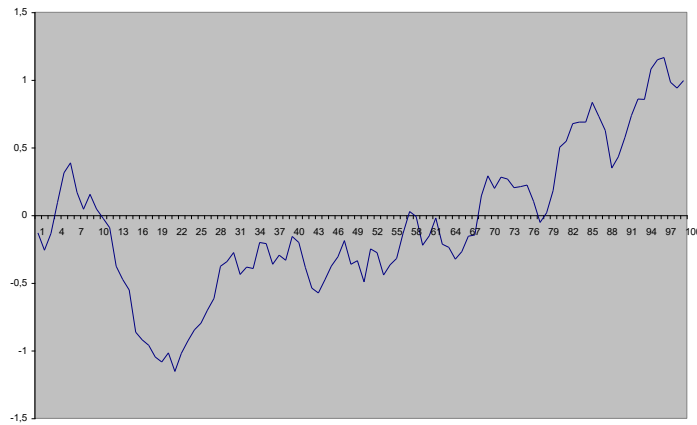


Рисунок 4. График компьютерной реализации последовательности H в случае модели Кокса-Росса-Рубинштейна с зависимыми значениями при $a = -0.1, b = 0.1, P(\delta_n = 1) = 0.6, \sigma = 0.1$.

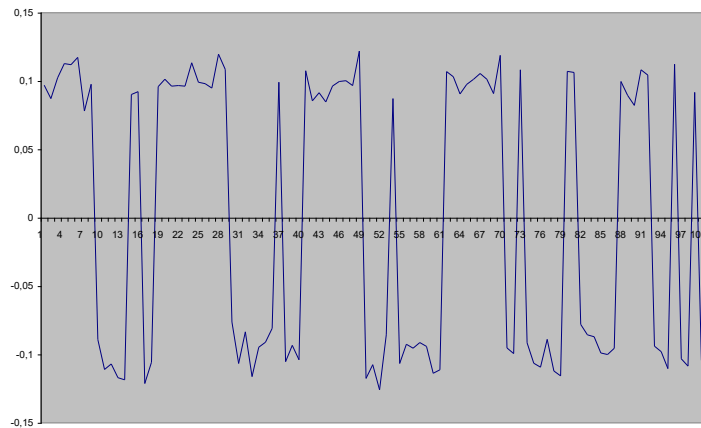


Рисунок 5. График машинной реализации последовательности h в случае модели Кокса-Росса-Рубинштейна с марковской цепью при

$$Q = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}, P(\delta_1 = 1) = 0.6, \sigma = 0.01.$$

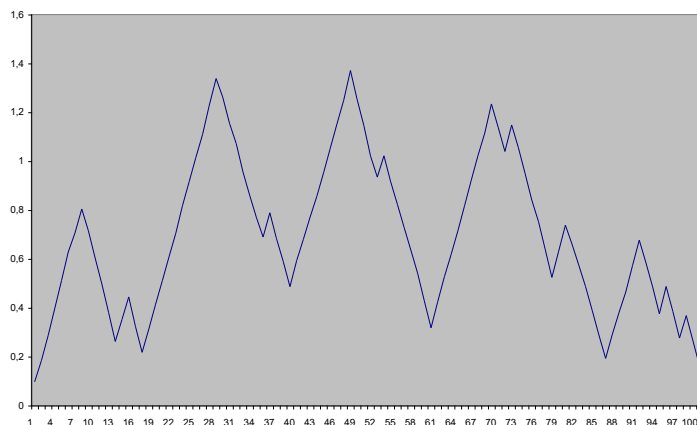


Рисунок 6. График машинной реализации последовательности H в случае модели Кокса-Росса-Рубинштейна с марковской цепью при

$$Q = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}, P(\delta_1 = 1) = 0.6, \sigma = 0.01.$$

Лекция 2

«Вероятностное описание временного ряда»

Спектральное представление стационарных в широком смысле
последовательностей

Определим функцию: $Z(\lambda) = \sum_{\{k: \lambda_k \leq \lambda\}} z_k$.

Использование функции $Z(\lambda)$ позволяет получить равенство:

$$h_n = \sum_{j=-\infty}^{\infty} \exp(i\lambda_j n) \Delta Z(\lambda_j), \text{ где } \Delta Z(\lambda_j) = Z(\lambda_j) - Z(\lambda_j -) = z_j.$$

Правая часть равенства напоминает интегральную сумму для интеграла $\int_{-\pi}^{\pi} \exp(i\lambda n) dZ(\lambda)$. Однако, функция $Z(\lambda)$ является случайной функцией. При фиксированном ω функция $Z(\lambda)$ может иметь неограниченную вариацию. Поэтому понимание интеграла для каждого ω как интеграла Римана-Стилтьеса неприемлемо.

Процессом с ортогональными приращениями назовем случайную функцию $Z(\omega, x)$, $x \in R$, если:

- 1) для каждого x $Z(\omega, x) \in H^2(\Omega, F, P)$;
- 2) для каждого $E|Z(\omega, x) - Z(\omega, x_n)|^2 \rightarrow 0$, $x_n \downarrow x$, $x_n \in R$;
- 3) для любых

$$x_1 < x_2 < x_3 < x_4 \quad E(Z(\omega, x_4) - Z(\omega, x_3)) \overline{(Z(\omega, x_2) - Z(\omega, x_1))} = 0.$$

Определим случайную меру полуоткрытого интервала $(a, b]$: $Z(\omega, (a, b]) = Z(\omega, b) - Z(\omega, a)$.

Определим случайную меру для суммы непересекающихся интервалов $\Delta = \sum_{k=1}^n (a_k, b_k]$: $Z(\omega, \Delta) = \sum_{k=1}^n (Z(\omega, b_k) - Z(\omega, a_k))$.

Рассмотрим алгебру \mathcal{A}_0 , порожденную полуоткрытыми интервалами $(a, b]$. Напомним, что элементарной стохастической мерой называется комплекснозначная функция $Z(\omega, \Delta)$, определенная для $\omega \in \Omega$ и $\Delta \in \mathcal{A}_0$, если:

- 1) $Z(\omega, \Delta) \in H^2(\Omega, F, P)$, $\forall \Delta \in \mathcal{A}_0$;
- 2) для любых непересекающихся $\Delta_1 \in \mathcal{A}_0$ и $\Delta_2 \in \mathcal{A}_0$ $Z(\omega, \Delta_1 + \Delta_2) = Z(\omega, \Delta_1) + Z(\omega, \Delta_2)$;
- 3) для любых непересекающихся $\Delta_1 \in \mathcal{A}_0, \Delta_2 \in \mathcal{A}_0, \dots$ таких, что

$$\Delta = \sum_{k=1}^{\infty} \Delta_k \in \mathcal{A}_0 \quad E \left| \Delta - \sum_{k=1}^n \Delta_k \right| \rightarrow 0, n \rightarrow \infty.$$

- 4) для любых непересекающихся Δ_1 и Δ_2 $E Z(\omega, \Delta_1) \overline{Z(\omega, \Delta_2)} = 0$.

Задача. Доказать выполнение 1)-4).

Определим функцию: $m(\Delta) = E|Z(\Delta)|^2$.

Как легко показать, $m(\Delta)$ является конечной мерой, и, следовательно, по теореме Каратеодори может быть продолжена на \mathcal{A} , где \mathcal{A} – минимальная σ -алгебра, содержащая \mathcal{A}_0 (борелевская σ -алгебра). Эту меру мы будем обозначать $m(\Delta)$ и называть структурной функцией элементарной стохастической меры $Z(\omega, \Delta)$.

Итак, перейдем к построению интеграла. Пусть $f(x) = \sum_{k=1}^n f_k I_{\Delta_k}(x)$,

$$f(x) = \sum_{k=1}^n f_k I_{\Delta_k}(x), \Delta_k \in \mathcal{A}_0.$$

Для каждой такой функции определим интеграл: $I(f) = \sum_{k=1}^n f_k Z(\omega, \Delta_k)$.

Отметим, если $g(x) = \sum_{k=1}^n g_k I_{\Delta_k}(x), \Delta_k \in \mathcal{A}_0$, то $E I(f) \overline{I(g)} = \sum_{k=1}^n f_k \bar{g}_k m(\Delta_k)$.

Пусть $L^2(R, \mathcal{A}, m)$ – гильбертово пространство комплекснозначных функций со скалярным произведением $\langle f, g \rangle = \int_R f(x) \bar{g}(x) m(dx)$. Таким образом, для любых функций заданного вида $(I(f), I(g)) = \langle f, g \rangle$. Это соотношение является ключевым при определении интеграла.

Пусть $f \in L^2$, тогда существует последовательность функций указанного вида таких, что $\|f - f_n\| \rightarrow 0$, при $n \rightarrow \infty$. Тогда $\|I(f_n) - I(f_m)\| = \|f_n - f_m\| \rightarrow 0$ при $m, n \rightarrow \infty$. То есть последовательность $I(f_n)$ является фундаментальной. Из полноты гильбертова пространства H^2 следует существование случайной величины $I(f) \in H^2$, что $I(f_n) \rightarrow I(f)$. Ее естественно назвать интегралом от функции $f \in L^2$ по стохастической мере Z . Естественно использовать для этой случайной величины интегральную запись $\int_R f(x) Z(dx)$.

Отметим основные свойства стохастического интеграла:

- 1) $(I(f), I(g)) = \langle f, g \rangle$;
- 2) $I(af + bg) = aI(f) + bI(g)$, где a и b – константы;
- 3) если $\|f_n - f\| \rightarrow 0$, то $\|I(f) - I(f_n)\| \rightarrow 0$.

Задача 2. Доказать свойства 1)-3).

Замечание. Поскольку $m(\Delta)$ конечна, то элементарная стохастическая мера может быть продолжена на σ -алгебру \mathcal{A} . Пусть $\Delta \in \mathcal{A}$, тогда $I_\Delta(x) \in L^2$. Определим: $\tilde{Z}(\omega, \Delta) = I(I_\Delta)$.

Задача 3. Доказать:

- 1) если $\Delta \in \mathcal{A}_0$, то $\tilde{Z}(\omega, \Delta) = Z(\omega, \Delta)$.
- 2) если $\Delta_1 \cap \Delta_2 = \emptyset$, то $\tilde{Z}(\omega, \Delta_1 + \Delta_2) = \tilde{Z}(\omega, \Delta_1) + \tilde{Z}(\omega, \Delta_2)$.
- 3) $E|\tilde{Z}(\omega, \Delta)|^2 = m(\Delta)$.
- 4) если $\Delta = \sum_{k=1}^{\infty} \Delta_k$, то $E\left|\tilde{Z}(\omega, \Delta) - \sum_{k=1}^n \tilde{Z}(\omega, \Delta_k)\right|^2 \rightarrow 0$ при $n \rightarrow \infty$.
- 5) если $\Delta_1 \cap \Delta_2 = \emptyset$, то $E\tilde{Z}(\omega, \Delta_1)\overline{\tilde{Z}(\omega, \Delta_2)} = 0$.

Пусть $R(n)$ – ковариационная функция стационарной в широком смысле последовательности $h = (h_n)$. Является справедливой теорема.

Теорема 1 (Герглотц). Существует конечная мера на $([-\pi, \pi), \mathcal{B}[-\pi, \pi))$ ($\mathcal{B}[-\pi, \pi)$ – борелевская σ -алгебра подмножеств интервала $[-\pi, \pi)$) $F(B)$, $B \in \mathcal{B}[-\pi, \pi)$, что для любого $n \in \mathbb{Z}$: $R(n) = \int_{-\pi}^{\pi} \exp(i\lambda n) F(d\lambda)$.

Обозначим через $F(\lambda) = F([-\pi, \lambda))$. В этом обозначении:

$$R(n) = \int_{-\pi}^{\pi} \exp(i\lambda n) dF(\lambda).$$

Если $F(\lambda)$ – абсолютно непрерывная функция: $F(\lambda) = \int_{-\pi}^{\lambda} f(x) dx$, то

$$R(n) = \int_{-\pi}^{\pi} \exp(i\lambda n) f(\lambda) d\lambda.$$

Следовательно, если $\sum_{n=-\infty}^{\infty} |R(n)|^2 < \infty$, то $f(\lambda) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} R(n) \exp(-i\lambda n)$.

Функция $F(\lambda)$ называется спектральной функцией, функция $f(\lambda)$ называется спектральной плотностью.

Если последовательность $h = (h_n)$ – вещественнозначная, то

$$R(n) = \frac{R(n) + R(-n)}{2} = \int_{-\pi}^{\pi} \cos \lambda n dF(\lambda),$$

если $F(\lambda)$ – абсолютно непрерывная функция, то

$$R(n) = \int_{-\pi}^{\pi} \cos \lambda n f(\lambda) d\lambda, f(\lambda) = \frac{1}{\pi} \left(\frac{R(0)}{2} + \sum_{n=1}^{\infty} R(n) \cos \lambda n \right).$$

Пусть $L^2(F) = L^2([-\pi, \pi), \mathcal{B}[-\pi, \pi), F)$ – гильбертово пространство комплекснозначных функций со скалярным произведением

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(\lambda) \overline{g(\lambda)} dF(\lambda).$$

Обозначим через $L_0^2(e)$ – линейное многообразие, порожденное функциями $e = (e_n(x) = \exp(ixn)), n \in \mathbf{Z}, x \in [-\pi, \pi)$. Отметим, что F – конечная мера на $\mathbf{B}[-\pi, \pi)$. Отсюда замыкание $L_0^2(e)$ совпадает с $L^2([-\pi, \pi), \mathbf{B}[-\pi, \pi), F)$.

Рассмотрим линейное многообразие $L_0^2(h)$, порожденное случайными величинами $h_n, n \in \mathbf{Z}$, и $L^2(h) = \overline{L_0^2(h)}$ (замыкание в среднеквадратичном смысле). Построим изоморфизм $\Phi: L^2(F) \Leftrightarrow L^2(h)$:

$$1) e_n \Leftrightarrow h_n.$$

2) $\sum a_n e_n \Leftrightarrow \sum a_n h_n, n \in \mathbf{Z}$ (предполагается, что только конечное число коэффициентов не равно нулю).

$$\text{Отметим, что } \langle e_n, e_m \rangle = \int_{-\pi}^{\pi} \exp(i(n-m)\lambda) dF(\lambda) = R(n-m) = \langle h_n, h_m \rangle,$$

$$\text{следовательно: } \langle \sum a_n e_n, \sum b_n e_n \rangle = \langle \sum a_n h_n, \sum b_n h_n \rangle.$$

$$\text{В частности, } \|\sum a_n e_n\|^2 = \|\sum a_n h_n\|^2 = \sum R(n-m) a_n \overline{a_n}.$$

Пусть $f \in L^2(F)$, тогда существует последовательность $f = (f_n) \in L_0^2(F): \lim_{n \rightarrow \infty} f_n = f$, откуда следует, что последовательность $\Phi(f_n) \in L_0^2(h)$ является фундаментальной последовательностью, следовательно, существует $g \in L^2(h): g = \lim_{n \rightarrow \infty} \Phi(f_n)$. Справедливо и обратное. Положим $f \Leftrightarrow g$, где f и g – рассмотренные выше элементы.

Задача 4.

Доказать, что построенное отображение $\Phi: L^2(F) \Leftrightarrow L^2(h)$ является изоморфизмом, сохраняющим скалярное произведение.

$$\text{Рассмотрим } f(\lambda) = I_{\Delta}(\lambda) \in L^2(f), \Delta \in \mathbf{B}([-\pi, \pi)).$$

Задача 5. Доказать, что процесс $Z(x) = \Phi(I_{[-\pi, x)}(\lambda))$ является процессом с ортогональными приращениями.

Пусть $f \in L_2(F)$. Покажем, что $I(f) = \Phi(f)$. Рассмотрим последовательность $\varphi = (\varphi_n)$ простых функций, сходящуюся к f , тогда последовательность $I(f_n)$ сходится к $I(f)$. С другой стороны, $I(f_n) = \Phi(f_n)$. Кроме этого, последовательность $\Phi(f_n)$ сходится к $\Phi(f)$. Следовательно, $I(f) = \Phi(f)$.

Возьмем функцию $e_n = \exp(i\lambda n)$, $I(e_n) = \int_{-\pi}^{\pi} \exp(i\lambda n) dZ(\lambda)$. С другой стороны, $\Phi(e_n) = h_n$. Следовательно, $h_n = \int_{-\pi}^{\pi} \exp(in\lambda) dZ(\lambda)$ – спектральное представление стационарной в широком смысле последовательности $h = (h_n)$.

Задача 6. Доказать, что $EZ(\lambda) = 0, E|Z(\lambda)|^2 = F(\lambda)$.

Пусть $h = (h_n)$ – стационарная в широком смысле последовательность, состоящая из действительных случайных величин. Тогда следует, что $h_n = \int_{-\pi}^{\pi} \exp(in\lambda) dZ(\lambda) = \bar{h}_n = \int_{-\pi}^{\pi} \exp(-in\lambda) d\bar{Z}(\lambda)$. Отсюда, $Z(\lambda) = \overline{Z(-\lambda)}$.

Задача 7. Представим $Z(\lambda)$ в виде $Z(\lambda) = Z_1(\lambda) + iZ_2(\lambda)$. Пусть $x_1 < x_2 < x_3 < x_4; x_1, x_2, x_3, x_4 \in [-\pi, \pi)$. Доказать, что $E(Z_1(x_2) - Z_1(x_1))(Z_1(x_3) - Z_1(x_4)) = E(Z_2(x_2) - Z_2(x_1))(Z_2(x_3) - Z_2(x_4)) = 0$.

Задача 8. Доказать: $h_n = \int_{-\pi}^{\pi} \cos \lambda n dZ_1(\lambda) - \int_{-\pi}^{\pi} \sin \lambda n dZ_2(\lambda)$.

Пусть $\xi \in L^2(h)$. Структура таких случайных величин определяется следующей теоремой.

Теорема 2. Если $\xi \in L^2(h)$, то найдется такая функция $\varphi \in L_2(F)$, что $\xi = \int_{-\pi}^{\pi} \varphi(\lambda) dZ(\lambda)$.

Доказательство. Поскольку $\xi \in L^2(h)$, то найдется последовательность

$g = (g_n)$ вида $g_n = \sum_{|k| \leq n} a_k h_k = \int_{-\pi}^{\pi} \left(\sum_{|k| \leq n} a_k \exp(i\lambda k) \right) dZ(\lambda)$, предел которой равен ξ .

Рассмотрим последовательность функций $\varphi_n = \sum_{|k| \leq n} a_k \exp(i\lambda k) = \Phi^{-1}(g_n)$.

Поскольку последовательность g_n – фундаментальна, то и последовательность φ_n – фундаментальна. Следовательно, существует предел $\lim_{n \rightarrow \infty} \varphi_n = \varphi$. Причем:

$$\varphi(\lambda) = \sum_{k=-\infty}^{\infty} a_k \exp(i\lambda k).$$

Линейный фильтр

Предположим, что на вход системы подается в момент времени m случайный сигнал x_m , при этом на выходе системы в момент времени n получается сигнал $f(n-m)x_m$, где $f(x)$, $x \in \mathbf{Z}$ – некоторая комплекснозначная функция, называемая импульсной переходной функцией.

Таким образом, суммарный сигнал на выходе устройства (фильтра):

$$y_n = \sum_{m=-\infty}^{\infty} f(n-m)x_m.$$

Для физически реализуемого устройства значение выходного сигнала зависит только от прошлых значений входного сигнала, то есть от значений x_m при $m \leq n$. Поэтому импульсная переходная функция физически реализуемого фильтра $f(x) = 0$ при $x < 0$. Следовательно, для физически реализуемого

фильтра:
$$y_n = \sum_{m=-\infty}^n f(n-m)x_m = \sum_{k=0}^{\infty} f(k)x_{n-k}.$$

Подадим на вход фильтра стационарную в широком смысле последовательность h с ковариационной функцией $R(n)$. На выходе получим случайную последовательность y . Вычислим

$$\text{Cov}(y_n, y_m) = \sum_{k, l=-\infty}^{\infty} f(k)R((n-m)+(l-k))\overline{f(l)}.$$

Сделаем замену переменных:

$$l-k=c, k=k. \text{ В результате: } \text{Cov}(y_n, y_m) = \sum_{k=-\infty}^{\infty} f(k) \left(\sum_{c=-\infty}^{\infty} R((n-m)+c)\overline{f(c+k)} \right).$$

В частности, $E|y_n|^2 = \sum_{k, l=-\infty}^{\infty} f(k)R(l-k)\overline{f(l)}$. Следовательно, для того,

чтобы $y_n \in L_2(h)$, необходимо и достаточно, чтобы $\sum_{k, l=-\infty}^{\infty} f(k)R(l-k)\overline{f(l)} < \infty$.

Ряд $\sum_{m=-\infty}^{\infty} f(n-m)h_m$ сходится в среднеквадратичном смысле, и,

следовательно, последовательность $y: y_n = \sum_{m=-\infty}^{\infty} f(n-m)h_m = \sum_{m=-\infty}^{\infty} f(m)h_{n-m}$ —

стационарная в широком смысле последовательность.

Из того, что $y_n \in L_2(h)$, следует, что для последовательности y справедливо спектральное представление со спектральной функцией:

$$\varphi(\lambda) = \exp(i\lambda n) \sum_{n=-\infty}^{\infty} f(m)\exp(-i\lambda m).$$

Функцию: $\psi(\lambda) = \sum_{n=-\infty}^{\infty} f(m)\exp(-i\lambda m)$ называют спектральной

характеристикой фильтра.

Задача 9. Доказать, что $\int_{-\pi}^{\pi} |\psi(\lambda)|^2 dF(\lambda) < \infty$.

Для ковариационной функции $R_y(x) = \sum_{k=-\infty}^{\infty} f(k) \left(\sum_{c=-\infty}^{\infty} R(x+c) \overline{f(c+k)} \right)$

справедливо представление: $R_y(n) = \int_{-\pi}^{\pi} |\psi(\lambda)|^2 \exp(i\lambda n) dF(\lambda)$.

В частности, если на вход фильтра подается белый шум $\varepsilon = (\varepsilon_n)$, то есть:

$E\varepsilon_n \varepsilon_m = \begin{cases} 1, n = m \\ 0, n \neq m \end{cases}$, то на выходе будет получаться стационарная

последовательность со спектральной плотностью: $f_y(\lambda) = \frac{1}{2\pi} |\psi(\lambda)|^2$.

Так как $f_y(\lambda)$ – спектральная плотность, то $\int_{-\pi}^{\pi} f_y(\lambda) d\lambda = 1 < \infty$.

Следовательно, $\int_{-\pi}^{\pi} |\psi(\lambda)|^2 d\lambda = 2\pi < \infty$. Поэтому функцию $\psi(\lambda)$ можно представить в виде ряда Фурье, с коэффициентами

$f(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\lambda m) \psi(\lambda) d\lambda$. В частности, если $f_y(\lambda) = \sum_{k=-\infty}^{\infty} a_k \exp(-i\lambda k)$, то

являются справедливыми равенства: $\frac{1}{2\pi} \sum_{m=-\infty}^{\infty} f(m) \overline{f(m-k)} = a_k$.

Из этих рассуждений вытекает следующая теорема.

Теорема 3. Если у стационарной в широком смысле последовательности существует плотность $f_y(\lambda)$, то последовательность можно представить в виде:

$$y_n = \sum_{m=-\infty}^{\infty} f(m) \varepsilon_{n-m}.$$

Доказательство. Пусть $y_n = \int_{-\pi}^{\pi} \exp(i\lambda n) dZ(\lambda)$. Рассмотрим процесс $\tilde{Z}(\lambda)$,

не зависящий от $Z(\lambda)$, с $E|\tilde{Z}(\lambda)|^2 = \frac{\lambda + \pi}{2\pi}$. Определим меру

$$\bar{Z}(\Delta) = \int_{\Delta} a(\lambda) dZ(\lambda) + \int_{\Delta} (1 - a(\lambda) \psi(\lambda)) d\tilde{Z}(\lambda), \quad \text{где } a(\lambda) = \begin{cases} \frac{1}{\psi(\lambda)}, \psi(\lambda) \neq 0 \\ 0, \psi(\lambda) = 0 \end{cases}, \quad \text{и}$$

процесс: $\bar{Z}(\lambda) = \int_{-\pi}^{\lambda} a(x) dZ(x) + \int_{-\pi}^{\lambda} (1 - a(x) \psi(x)) d\tilde{Z}(x)$.

Очевидно, что $E|\bar{Z}(\lambda)|^2 = \frac{\lambda + \pi}{2\pi}$.

Следовательно, $\varepsilon_n = \int_{-\pi}^{\pi} \exp(i\lambda n) d\bar{Z}(\lambda)$ является белым шумом. С другой

$$\text{стороны, } \int_{-\pi}^{\pi} \exp(i\lambda n) \psi(\lambda) d\bar{Z}(\lambda) = \int_{-\pi}^{\pi} \exp(i\lambda n) dZ(\lambda) = y_n.$$

Имеем:

$$\begin{aligned} \int_{-\pi}^{\pi} \exp(i\lambda n) \psi(\lambda) d\bar{Z}(\lambda) &= \int_{-\pi}^{\pi} \exp(i\lambda n) \sum_{m=-\infty}^{\infty} \exp(-i\lambda m) f(m) d\bar{Z}(\lambda) = \\ &= \sum_{m=-\infty}^{\infty} f(m) \int_{-\pi}^{\pi} \exp(i\lambda(n-m)) d\bar{Z}(\lambda) = \sum_{m=-\infty}^{\infty} f(m) \varepsilon_{n-m}. \end{aligned}$$

Замечание. Введение процесса $\tilde{Z}(\lambda)$ обусловлено возможностью равенства нулю спектральной плотности $f_y(\lambda)$. Если $f_y(\lambda) > 0$ (почти всюду по мере Лебега), то необходимость в процессе $\tilde{Z}(\lambda)$ отпадает.

$$\text{Пусть } \psi(\lambda) = \sum_{m=0}^{\infty} f(m) \exp(-i\lambda m), \quad \sum_{m=0}^{\infty} |f(m)|^2 < \infty, \quad \text{тогда}$$

$$\text{последовательность } y_n = \sum_{m=0}^{\infty} f(m) \varepsilon_{n-m}.$$

Статистическое оценивание параметров модели

Чтобы решать задачи анализа временных последовательностей, возникающих при описании эволюции стоимости финансовых активов, необходимо уметь оценивать по наблюдаемым значениям параметры модели.

Пусть h – стационарная в широком смысле последовательность с математическим ожиданием $Eh_n = m$ и ковариационной функцией

$$R(n) = \int_{-\pi}^{\pi} \exp(in\lambda) dF(\lambda). \text{ Возникает задача: как по наблюдениям } x_1, \dots, x_N$$

значений h_1, \dots, h_N получить «хорошие» оценки для m и $R(n)$.

$$\text{Рассмотрим в качестве оценки } m: m_N(x) = \frac{1}{N} \sum_{k=1}^N x_k.$$

Математическое ожидание $Em_N(h) = m$. Следовательно, оценка является несмещенной.

Дисперсия

оценки

$$Dm_N(h) = E\left(\frac{1}{N} \sum_{i=1}^N h_i - m\right)^2 = E\left(\frac{1}{N} \sum_{i=1}^N (h_i - m)\right)^2.$$

Отсюда

$$Dm_N(h) = \frac{1}{N^2} (NR(0) + 2(N-1)R(1) + \dots + 2R(N-1)) \leq \frac{2}{N} \sum_{i=0}^{N-1} R(i) - \frac{R(0)}{N}.$$

Используем спектральное представление ковариационной функции:

$$\frac{1}{N} \sum_{i=0}^{N-1} R(i) = \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\pi}^{\pi} \exp(i\lambda k) dF(\lambda) = \int_{-\pi}^{\pi} \frac{1}{N} \sum_{k=0}^{N-1} \exp(i\lambda k) dF(\lambda). \quad \text{Так как}$$

$$\frac{1}{N} \sum_{k=0}^{N-1} \exp(i\lambda k) = \begin{cases} 1, \lambda = 0 \\ \frac{1}{N} \frac{\exp(iN\lambda) - 1}{\exp(i\lambda) - 1}, \text{ то } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} R(i) = F(\{0\}). \text{ Если } F(\{0\}) = 0, \end{cases}$$

то $\lim_{N \rightarrow \infty} Dm_N(h) = 0$, и оценка $m_N(x)$ – состоятельна в среднем квадратическом смысле. Будем считать, что объем выборки наблюдений достаточно большой, чтобы оценка $m_N(x)$ хорошо приближала m . Поэтому дальше мы будем считать, что $m = 0$.

В качестве оценки ковариационной функции рассмотрим:

$$R_N(n, x) = \frac{1}{N-n} \sum_{k=1}^{N-n} x_k x_{n+k}.$$

Математическое ожидание $ER_N(n, h) = \frac{1}{N-n} \sum_{k=1}^{N-n} Eh_k h_{n+k} = R(n)$.

Следовательно, получена несмещенная оценка $R(n)$. Предположим, что для каждого n последовательность $\xi = (\xi_k)_{k \in \mathbb{Z}}$ с $\xi_k = h_k h_{k+n}$ является стационарной последовательностью. Рассмотрим дисперсию $DR_N(n, h)$. Аналогично предыдущему, для состоятельности оценки в среднеквадратическом смысле необходимым и достаточным условием является стремление к нулю $\frac{1}{N} \sum_{k=0}^{N-1} \hat{R}(k)$,

где $\hat{R}(k)$ – ковариационная функция последовательности ξ . Отметим, что $\hat{R}(k) = E\xi_k \xi_0 - R^2(n) = Eh_{n+k} h_k h_n h_0 - R^2(n)$. Если последовательность h является гауссовой последовательностью, то из $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} R^2(k) = 0$ следует

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \hat{R}(k) = 0.$$

Для демонстрации одного из самых эффективных методов метода максимального правдоподобия остановимся на оценке параметров модели авторегрессии q -го порядка: $h_n + \sum_{k=1}^q b_k h_k = \sigma \varepsilon_n$. Неизвестными параметрами модели являются b_1, \dots, b_q и σ . Предположим, что ε_n – независимые стандартные нормальные случайные величины. Обозначим через

$$y_n = x_n + \sum_{k=1}^q b_k x_{n-k}, \quad n = q+1, \dots, N. \text{ Логарифм функции правдоподобия:}$$

$$L(y_{q+1}, y_{q+2}, \dots, y_N) = - \left[\frac{N-q}{2} (\ln \sigma^2 + \ln 2\pi) + \frac{1}{2\sigma^2} \sum_{k=q+1}^N \left(x_k + \sum_{j=1}^q b_j x_{k-j} \right)^2 \right].$$

В соответствии с методом максимального правдоподобия, для получения оценок неизвестных параметров необходимо найти максимум

$L(y_{q+1}, y_{q+2}, \dots, y_N)$ или минимум $-L(y_{q+1}, y_{q+2}, \dots, y_N)$. Таким образом, оценки максимального правдоподобия являются решением системы линейных алгебраических уравнений:

$$R_{N-q}(r, x) + \sum_{j=1}^q R_{N-q}(r-j, x) b_j = 0,$$

$$\sigma_{N-q}^2 = \frac{1}{N-q} \sum_{j=q+1}^N \left(x_k + \sum_{j=1}^q b_j x_{k-j} \right)^2.$$

Пример.

Рассмотрим последовательность: $h_n = \sum_{j=-\infty}^{\infty} z_j \exp(i\lambda_j n)$,

где $Ez_j = 0$ и $E|z_j|^2 = \sigma_j^2 < \infty$, причем $\sum_{j=-\infty}^{\infty} \sigma_j^2 < \infty$; $-\pi \leq \lambda_1 < \lambda_2 < \dots < \lambda_n < \dots < \pi$.

Отметим, что ряд $\sum_{j=-\infty}^{\infty} z_j \exp(i\lambda_j n)$ сходится в среднем квадратическом и

$Eh_n = 0, R(k) = \sum_{j=-\infty}^{\infty} \sigma_j^2 \exp(i\lambda_j k)$. Следовательно, последовательность

стационарна.

Рассмотрим функцию $F(\lambda) = \sum_{\{j: \lambda_j \leq \lambda\}} \sigma_j^2$. $F(\lambda)$ – неубывающая кусочно-

постоянная функция, большая либо равная нулю, имеющая пределы слева и полунепрерывная справа. Следовательно, $F(\lambda)$ с точностью до нормировки является функцией распределения. Носитель этой функции сосредоточен на множестве $[-\pi, \pi)$. С помощью этой функции ковариационная функция рассматриваемой последовательности может быть записана в виде интеграла

$$\text{Лебега-Стилтьеса: } R(k) = \int_{-\pi}^{\pi} \exp(i\lambda k) dF(\lambda).$$

Пример.

Пусть ε – последовательность ортонормированных случайных величин с нулевыми математическими ожиданиями. Эту последовательность принято называть белым шумом. Понятно, что белый шум – стационарная

последовательность с ковариационной функцией $R(k) = \begin{cases} 1, k=0 \\ 0, k \neq 0 \end{cases}$. Эта функция

также может быть представлена в виде интеграла: $R(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\lambda k) d\lambda$.

Отсюда $dF(\lambda) = \frac{1}{2\pi} d\lambda$ или $F(\lambda) = \frac{\lambda + \pi}{2\pi}$, то есть является функцией распределения равномерного закона на интервале $[-\pi, \pi)$.

Лекция 3

«Автоковариационная и автокорреляционная функции. Декомпозиция временного ряда, тренд, сезонная и циклическая компоненты»

Линейные модели финансовых временных последовательностей

Определение.

Последовательность h называется последовательностью скользящего среднего, если она представима в виде: $h_n = \sum_{m=-\infty}^{\infty} f(m)\varepsilon_{n-m}$.

Если $f(m)=0$ при $m < 0$, то $h_n = \sum_{m=0}^{\infty} f(m)\varepsilon_{n-m}$, и последовательность называется односторонней последовательностью скользящего среднего. Если к тому же $f(m)=0$ при $m > p$, то $h_n = \sum_{m=0}^p f(m)\varepsilon_{n-m}$, и последовательность называется последовательностью скользящего среднего порядка p .

Спектральная плотность последовательности: $f_y(\lambda) = \frac{1}{2\pi} |P(\exp(-i\lambda))|^2$
с $P(z) = f(0) + f(1)z + \dots + f(p)z^p$.

Задача. Обосновать формулу для спектральной плотности последовательности.

Определение.

Последовательность $h_n + b_1 h_{n-1} + \dots + b_q h_{n-q} = \varepsilon_n$ называется последовательностью авторегрессии порядка q .

Найдем решение уравнения. Для этого определим оператор сдвига L : $Lx_n = x_{n-1}$.

Следовательно, $L^q x_n = x_{n-q}$. С использованием оператора L уравнение примет вид: $(1 + b_1 L + \dots + b_q L^q)h_n = \varepsilon_n$, где единица понимается как тождественный оператор.

Рассмотрим многочлен $P(z) = 1 + b_1 z + \dots + b_q z^q$. Тогда $P(z) = b_q (z - z_1) \dots (z - z_q) = (-1)^q b_q (z_1 - z) \dots (z_q - z)$, где z_k — корни многочлена $P(z)$. Отсюда уравнение приобретает вид: $(-1)^q b_q (z_1 - L) \dots (z_q - L)h_n = \varepsilon_n$, которое превращается в систему рекуррентных уравнений:

$$(z_1 - L)x_n^{(1)} = x_n^{(0)}$$

$$(z_2 - L)x_n^{(2)} = x_n^{(1)}$$

.....

$$(z_q - L)h_n = x_n^{(q-1)},$$

где $x_n^{(0)} = \frac{\varepsilon_n}{(-1)^q b_q}$. Рассмотрим k -е уравнение системы $(z_k - L)x_n^{(k)} = x_n^{(k-1)}$,

которое имеет вид: $x_n^{(k)} = \frac{1}{z_k} x_{n-1}^{(k)} + \frac{1}{z_k} x_n^{(k-1)}$.

Введем обозначение: $\alpha_k = \frac{1}{z_k}$. Решение уравнения:

$$x_n^{(k)} = \alpha_k^l x_{n-l}^{(k)} + \sum_{j=0}^{l-1} \alpha_k^{j+1} x_{n-j}^{(k-1)}.$$

Пусть $|\alpha_k| < 1$ и для любых k и n $E(x_n^{(k)})^2 < M$, тогда

$$E \left| x_n^{(k)} - \sum_{j=0}^{l-1} \alpha_k^{j+1} x_{n-j}^{(k-1)} \right|^2 = E \left| \alpha_k^l x_{n-l}^{(k-1)} \right|^2 = |\alpha_k|^{2l} E \left| x_{n-l}^{(k-1)} \right|^2 \rightarrow 0 \text{ при } l \rightarrow \infty. \text{ Тогда решение}$$

уравнения представляется в виде: $x_n^{(k)} = \sum_{j=0}^{\infty} \alpha_k^{j+1} x_{n-j}^{(k-1)}$.

Является справедливой теорема.

Теорема. Если корни многочлена $P(z)$ лежат вне единичной окружности, то решение уравнения является односторонним скользящим средним.

Доказательство. Для $k=1$ утверждение теоремы очевидно. Пусть оно справедливо для $k=r$, то есть $x_n^{(r)} = \sum_{j=0}^{\infty} a_j^{(r)} \varepsilon_{n-j}$, причем $\sum_{j=0}^{\infty} |a_j^{(r)}|^2 < \infty$.

Следовательно, $x_n^{(r+1)} = \sum_{k=0}^{\infty} \left(\sum_{l=0}^k \alpha_{r+1}^{l+1} a_{k-l}^{(r)} \right) \varepsilon_{n-k}$. Следовательно, $x_n^{(r+1)} = \sum_{k=0}^{\infty} a_k^{(r+1)} \varepsilon_{n-k}$,

где $a_k^{(r+1)} = \sum_{l=0}^k \alpha_{r+1}^{l+1} a_{k-l}^{(r)}$, причем $\sum_{j=0}^{\infty} |a_j^{(r+1)}|^2 < \infty$. Последнее справедливо потому,

что последовательность $a_j^{(r+1)}$ – свёртка последовательностей α_{r+1}^j и $a_j^{(r)}$, каждая

из которых принадлежит l_2 . Отсюда $h_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$.

Причем ряд, стоящий в правой части, сходится в среднеквадратическом смысле.

Спектральная плотность последовательности: $f_h(\lambda) = \frac{1}{2\pi |P(\exp(-i\lambda))|^2}$.

Задача. Доказать формулу спектральной плотности последовательности.

Пример.

Рассмотрим последовательность $h_n - \frac{5}{6} h_{n-1} + \frac{1}{6} h_{n-2} = \varepsilon_n$. Нули многочлена

$P(z) = 1 - \frac{5}{6}z + \frac{1}{6}z^2$: $z_1 = 2, z_2 = 3$ лежат вне единичной окружности.

Следовательно, последовательность h представима в виде одностороннего скользящего среднего. Рассмотрим $\frac{1}{P(z)} = \frac{A_1}{z - z_1} + \frac{A_2}{z - z_2}$, где

$$A_1 = \frac{1}{z_1 - z_2}, A_2 = \frac{1}{z_2 - z_1}. \text{ Тогда } h_n = P^{-1}(L)\varepsilon_n = 6\left(A_1(L - z_1)^{-1} + A_2(L - z_2)^{-1}\right)\varepsilon_n.$$

$$\text{Отсюда } h_n = 6\sum_{i=0}^{\infty} \left(\frac{A_1}{z_1^{i+1}} + \frac{A_2}{z_2^{i+1}}\right)\varepsilon_{n-i} = 6\sum_{i=0}^{\infty} \left(\frac{1}{3^{i+1}} - \frac{1}{2^{i+1}}\right)\varepsilon_{n-i}.$$

Определение. Модель случайной последовательности: $h_n + b_1 h_{n-1} + \dots + b_q h_{n-q} = a_0 \varepsilon_n + a_1 \varepsilon_{n-1} + \dots + a_p \varepsilon_{n-p}$ называется смешанной моделью авторегрессии и скользящего среднего.

Теорема. Если нули полинома $P(z) = 1 + b_1 z + \dots + b_q z^q$ лежат вне единичной окружности, то последовательность h представима в виде скользящего среднего и спектральная плотность последовательности h имеет

$$\text{вид: } f_h(\lambda) = \frac{1}{2\pi} \left| \frac{\widehat{P}(\exp(-i\lambda))}{P(\exp(-i\lambda))} \right|^2, \text{ где } \widehat{P}(z) = a_0 + a_1 z + \dots + a_p z^p.$$

Задача. Доказать теорему.

Линейный прогноз для стационарных в широком смысле последовательностей

Пусть h – стационарная в широком смысле последовательность с ковариационной функцией $R(n)$ и нулевым математическим ожиданием. Обозначим через $H_n = \overline{L^2}(h^n)$ и через $H = \overline{L^2}(h)$, где

$h^n = (\dots, h_{n-1}, h_n), h = (\dots, h_{n-1}, h_n, h_{n+1}, \dots)$. Пусть $H_{-\infty} = \bigcap_n H_n$. Для любого элемента

$\eta \in H$ обозначим через $\pi_n(\eta)$ проекцию η на подпространство H_n .

Соответственно $\pi_{-\infty}(\eta)$ – проекция η на $H_{-\infty}$. Каждый элемент

$\eta = \pi_{-\infty}(h) + (\eta - \pi_{-\infty}(h))$, причем $(\pi_{-\infty}(h), (\eta - \pi_{-\infty}(h))) = 0$. Таким образом,

$H = H_{-\infty} \otimes R$, где $H_{-\infty}$ состоит из элементов $\pi_{-\infty}(h)$, R из элементов $\eta - \pi_{-\infty}(h), \eta \in H$.

Определение. Стационарная последовательность называется регулярной, если: $H = R$.

Определение. Стационарная последовательность называется сингулярной, если: $H = H_{-\infty}$.

Теорема. Всякая стационарная в широком смысле последовательность представима в виде суммы: $h = \zeta + \xi$, где ξ – регулярная, ζ – сингулярная последовательности. При этом ζ и ξ ортогональны (ξ_n ортогональна ζ_m для любых n и m).

Действительно, для любого n $h_n \in H$. Причем $\zeta \in H_{-\infty}$, следовательно $H(\zeta) = H_{-\infty}$, $\xi \in R$, следовательно $H(\xi) = R$. Ортогональность ξ_n и ξ_m вытекает из ортогональности R и $H_{-\infty}$.

Задача. Доказать, что разложение, представленное в теореме, единственное.

Следующая теорема устанавливает связь между регулярной последовательностью и односторонним скользящим средним.

Теорема. Для того, чтобы невырожденная последовательность h была регулярной, необходимо и достаточно, чтобы она представлялась в виде последовательности одностороннего скользящего среднего.

Доказательство.

Необходимость. Представим $H_n = H_{n-1} \otimes B_n$. Поскольку H_n порождается элементами из H_{n-1} и элементами βh_n , и h – стационарная и регулярная, то размерность B_n равна единице. Пусть η_n – ненулевой элемент из B_n . Положим

$\varepsilon_n = \frac{\eta_n}{\|\eta_n\|}$. Теперь для любых фиксированных n и k представим:

$H_n = H_{n-k} \otimes B_{n-k+1} \otimes \dots \otimes B_n$. Отсюда $h_n = \sum_{i=0}^{k-1} a_i \varepsilon_{n-i} + \pi_{n-k}(h_n)$, так как $\varepsilon_{n-k+1}, \dots, \varepsilon_n$

образуют ортонормированный базис в $B_{n-k+1} \otimes \dots \otimes B_n$.

Рассмотрим предел $\lim_{i \rightarrow \infty} \sum_{i=0}^{k-1} a_i \varepsilon_{n-i}$. В силу неравенства

$\left\| \sum_{i=0}^{k-1} a_i \varepsilon_{n-i} \right\|^2 = \sum_{i=0}^{k-1} |a_i|^2 < \|h_n\|^2 < \infty$ ряд $\sum_{i=0}^{\infty} a_i \varepsilon_{n-i}$ сходится в среднеквадратическом

смысле. Покажем, что в том же самом смысле существует предел $\lim_{k \rightarrow \infty} \pi_{n-k}(h_n)$.

В силу стационарности достаточно рассмотреть случай $n=0$. Обозначим через

$\pi_i = \pi_i(h_0)$. Поскольку $\pi_{-k} = \pi_0 + \sum_{i=0}^k (\pi_{-i} - \pi_{-i+1})$ и слагаемые ортогональны,

имеем: $\sum_{i=0}^k \|\pi_{-i} - \pi_{-i+1}\|^2 = \left\| \sum_{i=0}^k (\pi_{-i} - \pi_{-i+1}) \right\|^2 = \|\pi_{-k} - \pi_0\|^2 \leq 4\|h_0\|^2$. Поэтому предел

существует. Для каждого k $\pi_{n-k} \in H_{n-k}$. Следовательно, предел $\lim_{k \rightarrow \infty} \pi_{n-k}(h_n)$

принадлежит $H_{-\infty}$. Так как $H_{-\infty}$ – тривиально по предположению, то

$\lim_{k \rightarrow \infty} \pi_{n-k}(h_n) = 0$.

Необходимость. Для любого n $H_n(h) \subseteq H_n(\varepsilon)$, $H_{-\infty}(h) \subseteq H_n(\varepsilon)$, ε_{n+1} ортогонально $H_n(\varepsilon)$, следовательно, ортогонально $H_{-\infty}(h)$. В то же время ε образуют базис в $H_n(h)$. Отсюда $H_{-\infty}(h)$ тривиально, а, значит, h – регулярна.

Из предыдущих теорем вытекает разложение Вольда.

Теорема. Если h – невырожденная последовательность, то:

$$h_n = \zeta_n + \sum_{i=0}^{\infty} a_i \varepsilon_{n-i}.$$

Откуда можно получить наилучший линейный прогноз для h_n . Как было показано ранее, наилучшим линейным прогнозом для h_n по прошлым значениям h^0 является проекция $\hat{h}_n = \pi_{L^2(h^0)}(h_n)$. Из разложения Вольда следует,

$$\text{что } \hat{h}_n = \pi_{L^2(\zeta^0)}(h_n) + \pi_{L^2(\varepsilon^0)}(h_n). \text{ Отсюда } \hat{h}_n = \zeta_n + \sum_{i=n}^{\infty} a_i \varepsilon_{n-i}.$$

$$\text{Дисперсия ошибки равна: } \sigma_n^2 = \sum_{i=0}^{n-1} |a_i|^2.$$

Если последовательность h – сингулярная, то $\hat{h}_n = \zeta_n$ и дисперсия ошибки равна нулю (возможно точное предсказание), если h – регулярная, то

$$\hat{h}_n = \sum_{i=n}^{\infty} a_i \varepsilon_{n-i} \text{ и } \sigma_n \leq \sigma_{n+1}. \text{ Причем в случае регулярной последовательности } \lim_{n \rightarrow \infty} \sigma_n^2 = E|h_n|^2, \text{ то есть } \lim_{n \rightarrow \infty} \hat{h}_n = 0.$$

Интерес представляет прогноз для регулярной последовательности.

Формула $\hat{h}_n = \sum_{i=n}^{\infty} a_i \varepsilon_{n-i}$ даёт решение задачи оптимального линейного прогноза, однако, для того, чтобы ей можно было бы воспользоваться, необходимо решить следующую задачу. Задана ковариационная функция $R(n)$ или спектральная плотность $f(\lambda)$. Требуется вычислить коэффициенты a_k и выразить ε_n через \dots, h_{n-1}, h_n .

Рассмотрим решение этой задачи. Как было показано выше, если последовательность представима в виде скользящего среднего, то для такой последовательности существует спектральная плотность $f(\lambda) = \frac{1}{2\pi} |\varphi(\lambda)|^2$.

Обратно, если $f(\lambda) > 0$, то существует представление последовательности в виде скользящего среднего. Ограничимся рассмотрением важного для приложений случая: $\varphi(\lambda) = \Phi(\exp(-i\lambda))$, где $\Phi(z) = \sum_{k=0}^{\infty} b_k z^k$, радиус сходимости степенного ряда $r > 1$ и нули $\Phi(z)$ лежат вне единичной окружности.

Рассмотрим спектральное представление h_n и \hat{h}_n : $h_n = \int_{-\pi}^{\pi} \exp(i\lambda n) dZ(\lambda)$,

$$\hat{h}_n = \int_{-\pi}^{\pi} \varphi_n(\lambda) dZ(\lambda).$$

$$\text{Дисперсия ошибки: } E|h_n - \hat{h}_n|^2 = \int_{-\pi}^{\pi} |\exp(i\lambda n) - \varphi_n(\lambda)|^2 f(\lambda) d\lambda.$$

Таким образом, требуется найти минимум правой части равенства по $\varphi_n \in H_0(F)$. Покажем, что функция $\hat{\varphi}_n(\lambda) = \exp(i\lambda n) \frac{\Phi_n(\lambda)}{\Phi(\lambda)}$ доставляет минимум правой части равенства.

Во-первых, $\hat{\varphi}_n \in H_0(F)$. Во-вторых, для любого $m \geq 0$ скалярное произведение $(\exp(i\lambda n) - \hat{\varphi}_n(\lambda), \exp(-i\lambda m)) = 0$. Действительно, $(\exp(i\lambda n) - \hat{\varphi}_n(\lambda), \exp(-i\lambda m)) =$

$$\begin{aligned} & \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\lambda(m+n)) [\Phi(\exp(-i\lambda)) - \Phi_n(\exp(-i\lambda))] \overline{\Phi(\exp(-i\lambda))} d\lambda = \\ & \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\lambda m) \left[\sum_{k=0}^{n-1} b_k \exp(i\lambda(n-k)) \right] \left[\sum_{l=0}^{\infty} \overline{b_k} \exp(i\lambda l) \right] d\lambda = 0, \quad \text{так как} \\ & \int_{-\pi}^{\pi} \exp(i\lambda k) d\lambda = 0, \quad \text{для любого } k > 1. \end{aligned}$$

Разлагая функцию $\hat{\varphi}_n(\lambda)$ в ряд Фурье: $\hat{\varphi}_n(\lambda) = \sum_{i=0}^{\infty} c_{-i} \exp(-i\lambda)$, получим оптимальный линейный прогноз: $\hat{h}_n = \sum_{i=0}^{\infty} c_{-i} h_{n-i}$.

Подчеркнем, что из разложения Вольда регулярной последовательности $h_n = \sum_{i=0}^{\infty} a_k \varepsilon_{n-k}$ следует, что спектральная плотность допускает представление:

$$f(\lambda) = \frac{1}{2\pi} |\Phi(\exp(-i\lambda))|^2, \quad \text{где } \Phi(z) = \sum_{k=0}^{\infty} a_k z^k. \quad \text{Обратно, если спектральная}$$

плотность допускает представление $f(\lambda) = \frac{1}{2\pi} |\Phi(\exp(-i\lambda))|^2$, где

$$\Phi(z) = \sum_{k=0}^{\infty} a_k z^k, \quad \text{то разложение Вольда регулярной последовательности}$$

$h_n = \sum_{i=0}^{\infty} a_k \varepsilon_{n-k}$. Таким образом, задача представления спектральной плотности и задача разложения регулярной последовательности эквивалентны.

Задача. Доказать, что $\exp(i\lambda n) \frac{\Phi_n(\lambda)}{\Phi(\lambda)} \in H_0(F)$.

Спектральная плотность последовательности $h_n - \alpha h_{n-1} = \frac{1}{2}(2\varepsilon_n + \varepsilon_{n-1})$:

$$f(\lambda) = \frac{1}{2\pi} \left| \frac{P(\exp(-i\lambda))}{Q(\exp(-i\lambda))} \right|, \quad \text{где} \quad P(z) = \frac{z+2}{2}, Q(z) = 1 - \alpha z. \quad \text{Функция}$$

$$\Phi(z) = 1 + \sum_{k=1}^{\infty} \frac{\alpha^{k-1}(1+2\alpha)z^k}{2}. \quad \text{Отсюда} \quad \hat{\varphi}_1(\lambda) = \frac{1+2\alpha}{\exp(-i\lambda)+2}, \quad \text{следовательно,}$$

$$\text{прогноз на один шаг: } h_1 = \frac{1}{2}(1+2\alpha) \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} h_{-k}.$$

Безгранично делимые распределения

Перейдем к рассмотрению класса распределений, который включает в себя устойчивые распределения.

Определение.

Распределение F называется безгранично делимым, если для каждого n существует такое распределение F_n , что $F = F_n^{(n)}$.

Другими словами, для каждого n существуют n независимых случайных величин с $S_n = Y_1^n, \dots, Y_n^n$ и одним и тем же распределением F_n таким, что сумма $S_n = Y_1^n + \dots + Y_n^n$ имеет распределение F .

Как уже отмечалось выше, класс устойчивых распределений включен в класс безгранично делимых распределений. Действительно, пусть X_1, \dots, X_n, \dots — независимые копии X и распределение F_X устойчиво с индексом устойчивости α .

$$\text{Определим: } Y_k^n = \frac{X_k}{n^{1/\alpha}} - \frac{D_n}{n^{1/\alpha+1}}.$$

Тогда сумма $Y_1^n + \dots + Y_n^n$ имеет распределение F .

В качестве первого примера рассмотрим распределение Пуассона. Характеристическая функция распределения Пуассона, как уже показывалось ранее, $\varphi_F(\theta) = \exp(\lambda(\exp(i\theta) - 1))$. Поскольку характеристическая функция

свертки $\varphi_{F^{(n)}}(\theta) = (\varphi_F(\theta))^n$, то имеем распределение Пуассона с параметром

$$\lambda_n = \frac{\lambda}{n}. \quad \text{Следовательно, распределение Пуассона безгранично делимое}$$

распределение.

Рассмотрим ν -гамма-распределение с плотностью:

$$f_{\lambda, \nu}(x) = \begin{cases} \frac{1}{\Gamma(\nu)} \lambda^\nu x^{\nu-1} \exp(-\lambda x), & x \geq 0 \\ 0, & x < 0 \end{cases},$$

где $\lambda > 0, \nu > 0$ — параметры распределения. В частности, при $\nu = 1$ получается

$$\text{показательное распределение: } f_{\lambda, 1}(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Рассмотрим свертку гамма-распределений:

$$f_{\lambda,\nu} * f_{\lambda,\mu} = \frac{\lambda^{\nu+\mu}}{\Gamma(\nu)\Gamma(\mu)} \exp(-\lambda x) \int_0^x (x-y)^{\nu-1} y^{\mu-1} dy.$$

Сделаем подстановку в интеграле $u = xt$, в результате получим: $\int_0^x (x-y)^{\nu-1} y^{\mu-1} dy = x^{\nu+\mu-1} \int_0^1 (1-t)^{\nu-1} t^{\mu-1} dt$.

Интеграл в этой формуле – бета-функция: $B(\nu, \mu) = \int_0^1 (1-t)^{\nu-1} t^{\mu-1} dt = \frac{\Gamma(\nu)\Gamma(\mu)}{\Gamma(\nu+\mu)}$.

В результате: $f_{\lambda,\nu} * f_{\lambda,\mu} = \frac{\lambda^{\nu+\mu}}{\Gamma(\nu+\mu)} x^{\nu+\mu-1} \exp(-\lambda x) = f_{\lambda,\nu+\mu}$.

Замкнутость свертки для гамма-распределений позволяет утверждать, что гамма-распределение безгранично делимо. Действительно для гамма-распределения $f_{\lambda,\mu} - f_{\lambda,\mu/n}^{(n)} = f_{\lambda,\mu}$. В частности, для показательного распределения $f_{\lambda,1/n}^{(n)} = f_{\lambda,1}$.

Рассмотрим вырожденную случайную величину, сидящую в одной точке: $P(X=a)=1$. Характеристическая функция распределения: $\varphi_a(\theta) = \exp(i\theta a)$. Поскольку $\varphi_a^n(\theta) = \exp(i\theta na)$, то Y_k^n – вырожденные случайные величины, сидящие в точке $\frac{a}{n}$.

В результате работ Б. де Финнети, А.Н. Колмогорова, П. Леви и А.Я. Хинчина получена формула для характеристических функций безгранично делимых распределений:

$$\varphi_X(\theta) = \exp \left\{ ib\theta - c\theta^2 + \int_{-\infty}^{\infty} (\exp(i\theta x) - 1 - i\theta x I(|x| \leq 1)) \nu(dx) \right\},$$

где $c \geq 0$, $\nu = \nu(dx)$ – положительная мера (называемая мерой Леви), удовлетворяющая условиям:

$$\nu(\{0\}) = 0,$$

$$\int_{-\infty}^{\infty} (x^2 \wedge 1) \nu(dx) < \infty.$$

Таким образом, безгранично делимое распределение описывается триплетом (b, c, ν) , который определяется единственным образом.

В качестве примера вернемся к распределению Пуассона с характеристической функцией $\varphi_X(\theta) = \exp\{\lambda(\exp(i\theta) - 1)\}$ для которой триплет $(0, 0, \nu(dx) = \lambda I_{\{1\}} dx)$.

Приведём графики последовательностей скользящего среднего, авторегрессии и смешанной модели. Приведём пример построения прогноза для конкретной модели.

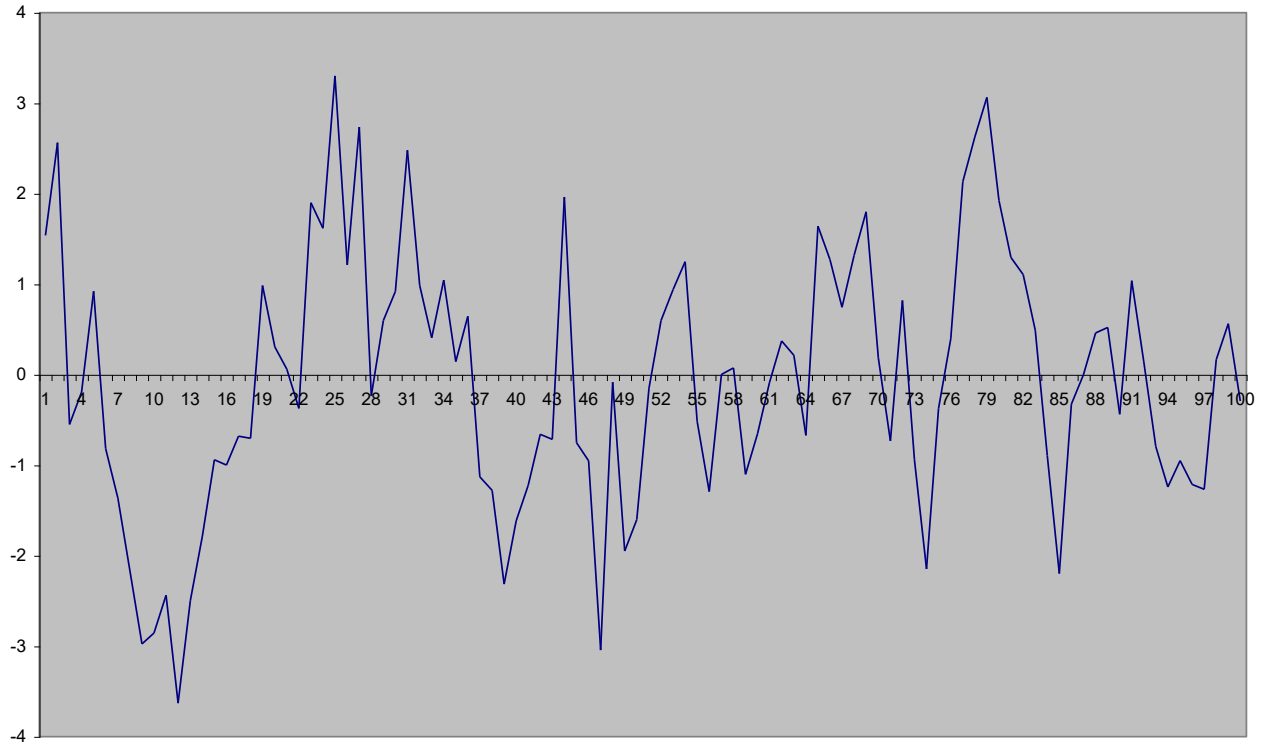


Рисунок 1. Машинная реализация последовательности скользящего среднего

порядка $p = 4$ с коэффициентами $f(i) = \frac{1}{i+1}, i = 0, 1, \dots, 4$.

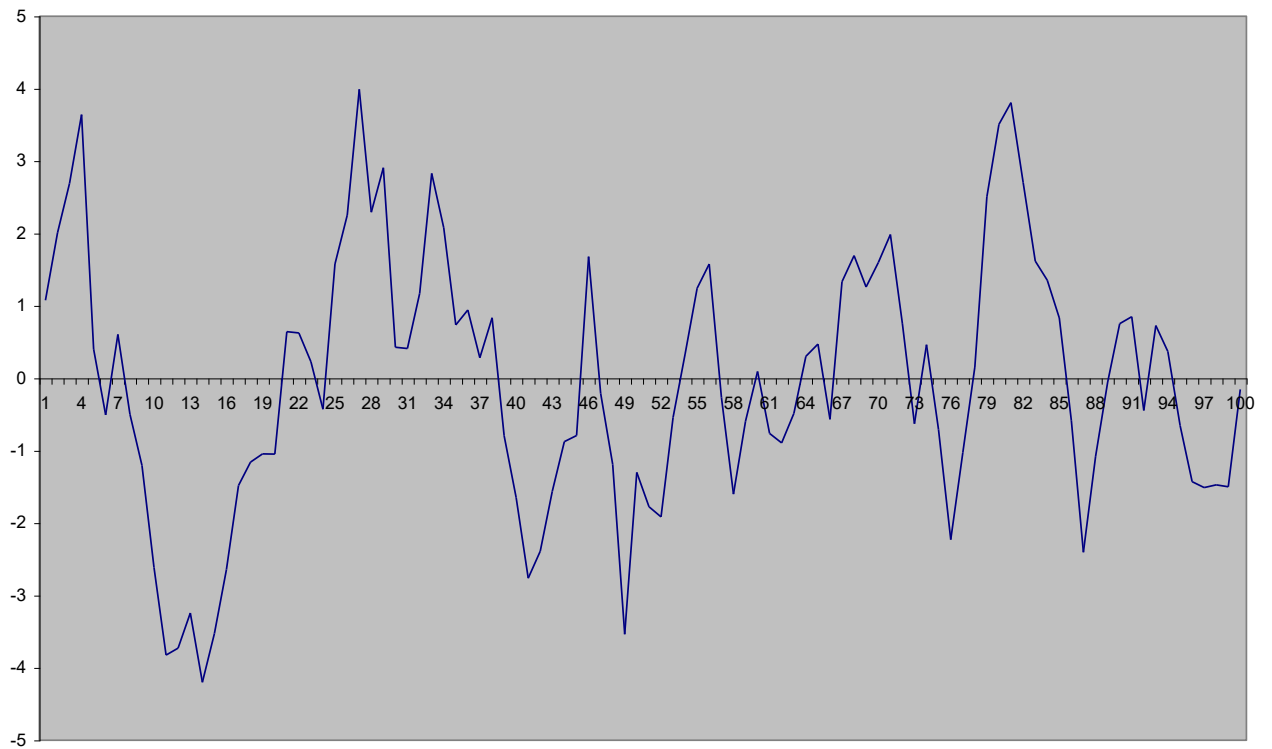


Рисунок 2. Машинная реализация авторегрессионной последовательности:

$$h_n - \frac{5}{6}h_{n-1} + \frac{1}{6}h_{n-2} = \varepsilon_n.$$

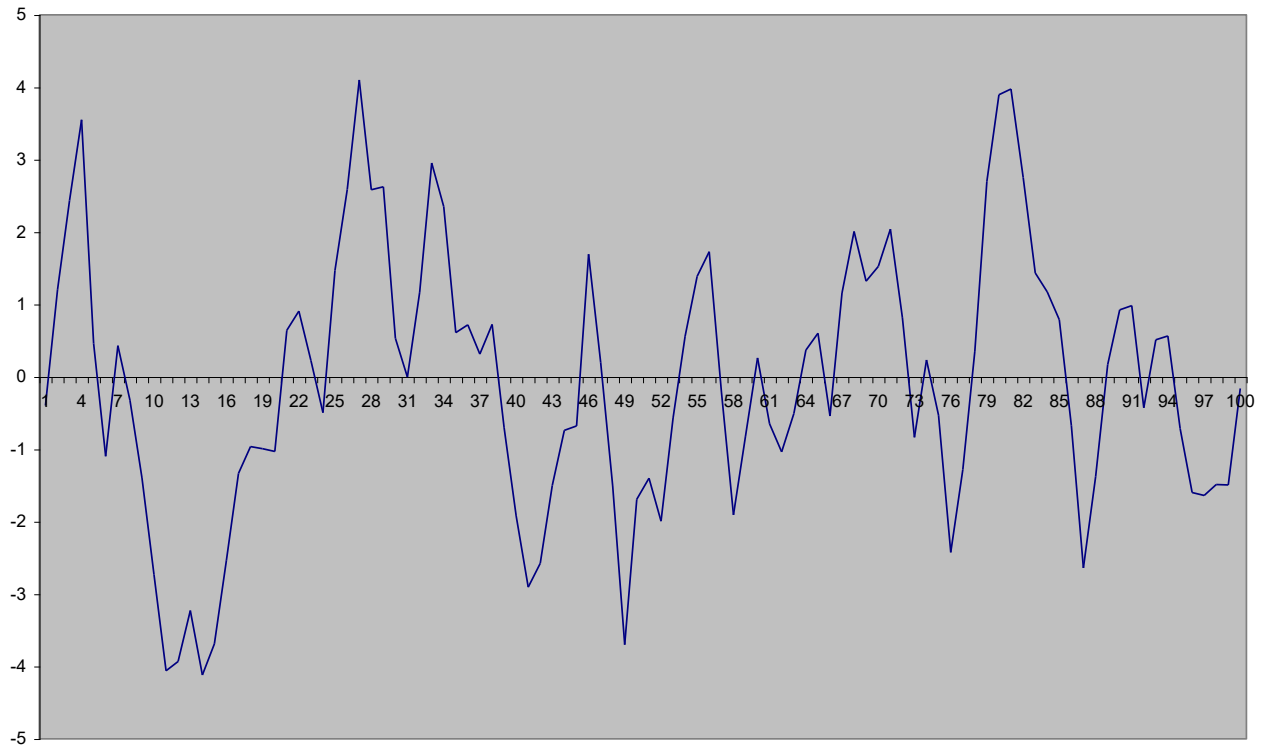


Рисунок 3. График машинной реализации последовательности

$$h_n - \frac{h_{n-1}}{2} = \frac{2\varepsilon_n + \varepsilon_{n-1}}{2}.$$

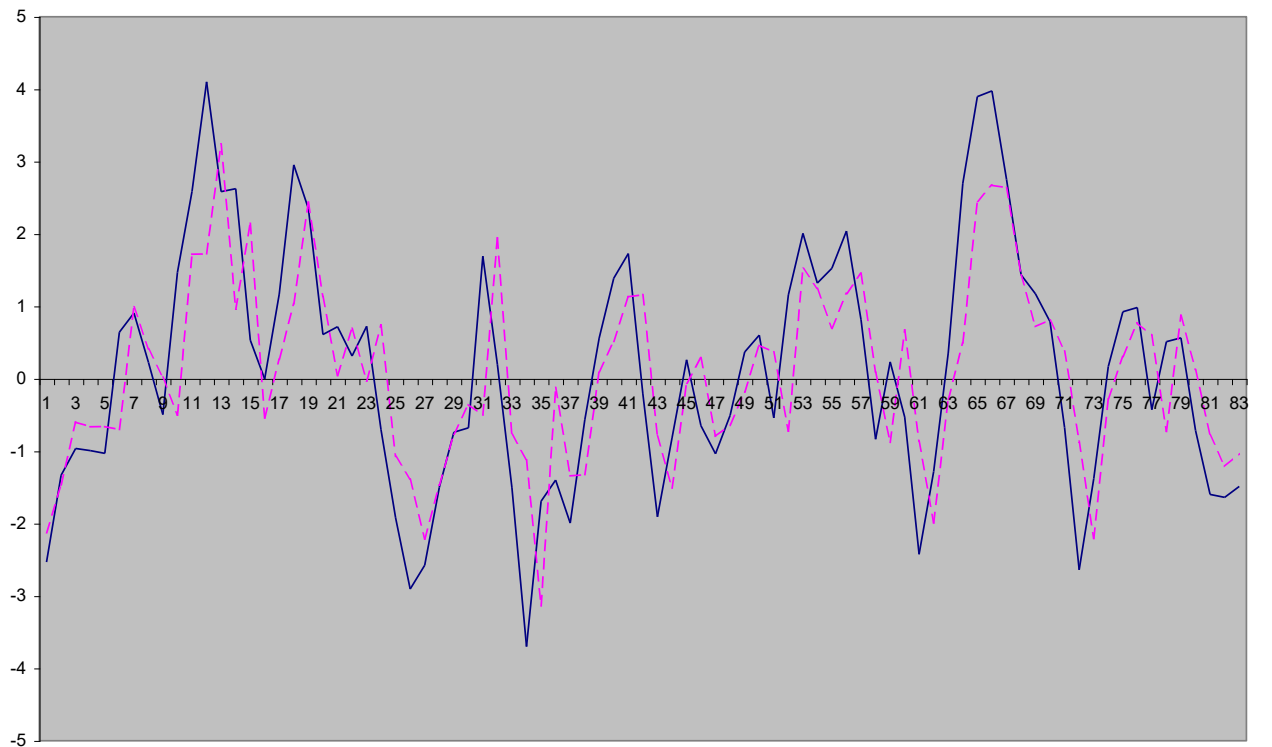


Рисунок 4. График машинной реализации последовательности

$h_n - \frac{h_{n-1}}{2} = \frac{2\varepsilon_n + \varepsilon_{n-1}}{2}$ (сплошная линия) и последовательности \hat{h} (пунктирная линия).

Лекция 4

«Элементы кластерного анализа. Метод k-средних»

Рассмотрим выборку значений возврата $R = \{\rho_i\}_{i=1}^N$, состоящую из независимых и одинаково распределенных случайных величин с плотностью общего закона распределения: $p(x) = p_1 p(x/q_1) + p_2 p(x/q_2)$. Здесь q_1 и q_2 – наборы параметров распределений, подлежащих оценке.

Необходимо выделить два непересекающихся интервала минимальной длины $[a_1, a_2]$ и $[a_3, a_4]$, вероятность попадания в которые была бы не меньше заданной доверительной вероятности. То есть, необходимо рассмотреть задачу оценки параметров модели, а именно, a_1, a_2, a_3 и a_4 . Для этого рассмотрим параметрические и непараметрические методы. В качестве параметрического метода рассмотрим метод максимального правдоподобия. В качестве непараметрического метода рассмотрим метод k-средних. Рассмотрим приложение данной задачи в стохастической финансовой математике, а именно, рассмотрим обобщённую модель Кокса-Росса-Рубинштена, в которой параметры модели заменяются на непересекающиеся интервалы. Эта робастная постановка задачи значительно обобщает модель и делает её более приближенной к моделям на реальном финансовом рынке. Приводится алгоритм расчёта интервала справедливых цен для Европейского и Азиатского опционов.

Параметрические методы.

Максимальное правдоподобие. Оценки по методу максимального правдоподобия являются решением оптимизационной задачи:

$$\max_{p_1, p_2, q_1, q_2} \left[\sum_{i=1}^N \ln(p_1 p(\rho_i/q_1) + p_2 p(\rho_i/q_2)) \right].$$

Максимум ищется по параметрам законов распределений и вероятностям p_1 и p_2 . Следующий алгоритм является монотонным и сходящимся к множеству стационарных точек, среди которых есть решение. Для нормальных законов алгоритм сходится к максимально правдоподобным оценкам.

Алгоритм:

1. Определяются p_1 и p_2 ; параметры распределений: q_1 и q_2 .

$$\alpha_{i,1} = \frac{p(r_i/q_1)p_1}{p(r_i/q_1)p_1 + p(r_i/q_2)p_2}, \alpha_{i,2} = \frac{p(r_i/q_2)p_2}{p(r_i/q_1)p_1 + p(r_i/q_2)p_2},$$

2. $p_1 = \frac{1}{N} \sum_{i=1}^N \alpha_{i,1}, p_2 = \frac{1}{N} \sum_{i=1}^N \alpha_{i,2},$

$$q_1 = \arg \max \sum_{i=1}^N \alpha_{i,1} \ln p_1(\rho_i/q), q_2 = \arg \max \sum_{i=1}^N \alpha_{i,2} \ln p_2(\rho_i/q)$$

3. Если «критерий остановки» то стоп иначе 2.

Алгоритм является монотонным и сходящимся к множеству стационарных точек, среди которых есть решение. Для нормальных законов алгоритм сходится к максимально правдоподобным оценкам.

Нормальный закон. Для нормальных законов параметры: $q_1 = \langle m_1, \sigma_1^2 \rangle, q_2 = \langle m_2, \sigma_2^2 \rangle$. Очередное приближение к максимально правдоподобным оценкам вычисляется следующим образом:

$$\alpha_{i,1} = \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\rho_i - m_1)^2}{2\sigma_1^2}\right) p_1}{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\rho_i - m_1)^2}{2\sigma_1^2}\right) p_1 + \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\rho_i - m_2)^2}{2\sigma_2^2}\right) p_2},$$

$$\alpha_{i,2} = \frac{\frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\rho_i - m_2)^2}{2\sigma_2^2}\right) p_2}{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\rho_i - m_1)^2}{2\sigma_1^2}\right) p_1 + \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\rho_i - m_2)^2}{2\sigma_2^2}\right) p_2}$$

$$p_1 = \frac{1}{N} \sum_{i=1}^N \alpha_{i,1}, p_2 = \frac{1}{N} \sum_{i=1}^N \alpha_{i,2},$$

$$m_1 = \frac{1}{\sum_{i=1}^N \alpha_{i,1}} \sum_{i=1}^N \alpha_{i,1} \rho_i, m_2 = \frac{1}{\sum_{i=1}^N \alpha_{i,2}} \sum_{i=1}^N \alpha_{i,2} \rho_i,$$

$$\sigma_1^2 = \frac{1}{\sum_{i=1}^N \alpha_{i,1}} \sum_{i=1}^N \alpha_{i,1} \rho_i^2 - m_1^2, \sigma_2^2 = \frac{1}{\sum_{i=1}^N \alpha_{i,2}} \sum_{i=1}^N \alpha_{i,2} \rho_i^2 - m_2^2$$

Как уже отмечалось, для двух нормальных законов алгоритм сходится к максимально правдоподобным оценкам. Для выбора начальных значений для p_1, p_2 и q_1, q_2 выборка разбивается на два подмножества медианой выборки:

$R_1 = \{\rho \in R, \rho \leq \text{med}(R)\}, R_2 = R \setminus R_1$, после этого вычисляются $p_1 = \frac{|R_1|}{|R|}, p_2 = \frac{|R_2|}{|R|}$,

$$m_1 = \frac{1}{|R_1|} \sum_{\rho \in R_1} \rho, m_2 = \frac{1}{|R_2|} \sum_{\rho \in R_2} \rho, \sigma_1^2 = \frac{1}{|R_1|} \sum_{\rho \in R_1} \rho^2 - m_1^2, \sigma_2^2 = \frac{1}{|R_2|} \sum_{\rho \in R_2} \rho^2 - m_2^2.$$

После остановки разбиение выборки R на два класса R_1 и R_2 осуществляется следующим образом. Элемент выборки ρ_i относится к R_1 , если $p(\rho_i / q_1) p_1 \geq p(\rho_i / q_2) p_2$, иначе ρ_i относится к R_2 .

Непараметрические методы. К непараметрическим методам относятся методы k -средних и нечетких k -средних, а также метод иерархической кластеризации. Остановимся на методе k -средних. В основе метода k -средних лежит оптимизационная задача:

$$\min_{m,x} \sum_{i=1}^N (\rho_i - m_1)^2 x_i + \sum_{i=1}^N (\rho_i - m_2)^2 (1 - x_i), x_i \in \{0,1\}.$$

Применение координатного спуска к этой задаче – основной способ кластеризации методом k -средних. Хорошо известно, что, являясь монотонным, метод не гарантирует сходимости к решению оптимизационной задачи. Существенно упрощающей особенностью рассматриваемой задачи

является ее одномерность. Одномерность данных позволяет перед кластеризацией упорядочить данные в порядке возрастания. При условии, что число кластеров равно двум, можно предложить следующий алгоритм.

Алгоритм метода k - средних.

1. Определяем $R_1 = \{\rho_1\}, R_2 = \{\rho_2, \rho_2, \dots, \rho_N\}$.
2. Вычисляем $m_1 = \frac{1}{|R_1|} \sum_{r \in R_1} r, m_2 = \frac{1}{|R_2|} \sum_{r \in R_2} r$.
3. Вычисляем следующее разбиение. Допустим $R_1 = \{\rho_1, \rho_1, \dots, \rho_k\}$, тогда следующее разбиение $R_1 = \begin{cases} R_1 + \{\rho_{k+1}\}, d(\rho_{k+1}, m_1) < d(\rho_{k+1}, m_2) \\ R_1, \text{ иначе} \end{cases}, R_2 = R \setminus R_1$.
4. Если «критерий остановки» то стоп иначе 2.

После кластеризации. Далее мы используем важное в статистике определение квантиля. Под α -квантилем функции распределения вероятностей $F(x)$ понимается решение уравнения $F(x) = \alpha$, которое не всегда существует, если функция распределения разрывна. Поэтому более естественным является следующее определение: верхним квантилем называется $x_\alpha^+ = \min\{x : F(x) \geq \alpha\}$, нижним квантилем называется $x_\alpha^- = \max\{x : F(x) \leq \alpha\}$. Мы будем использовать это понятие при вычислении параметров модели a_1, a_2, a_3 и a_4 . Для вычисления параметров a_1 и a_2 используется первый кластер $R_1 = \{\rho_1, \rho_2, \dots, \rho_l\}$ для вычисления параметров a_3 и a_4 – второй кластер $R_2 = \{\rho_{l+1}, \rho_{l+2}, \dots, \rho_N\}$. Вместо неизвестной функции распределения будем использовать эмпирическую функцию распределения $F_s^1(x)$, соответствующей выборке R_1 , с помощью которой вычислим вероятность попадания случайной величины в интервал - $P(x < \xi \leq y) = F_e^1(y) - F_e^1(x)$. Задача вычисления параметров a_1 и a_2 по выборке R_1 заключается в вычислении $\min(y - x)$ при ограничении $F_e^1(y) - F_e^1(x) \geq \alpha$. Для решения задачи рассмотрим семейство μ -задач: $\max x, F_e^1(x) \leq \mu(1 - \alpha); \min y, F_e^1(y) \geq \alpha(1 - \mu) + \mu$. Параметр $\mu \in [0, 1]$. Согласно определению решением μ -задачи является интервал $(x_{\mu(1-\alpha)}^-, x_{(1-\mu)\alpha+\mu}^+)$. Далее вычисляется $\min(x_{(1-\mu)\alpha+\mu}^+ - x_{\mu(1-\alpha)}^-)$, при этом рассматривается приближение задачи, а именно, $\mu \in \{\mu_0, \mu_2, \dots, \mu_s\}$, где $(\mu_i)_{i=0}^s$ - разбиение интервала $[0, 1]$. После решения задачи получаем $a_1 = x_{\mu^*(1-\alpha)}^-, a_2 = x_{(1-\mu^*)\alpha+\mu^*}^+$. Аналогично, вычисляются оценки a_3 и a_4 .

Предварительная обработка выборки. В результате анализа реальных данных, особенно данных, полученных с малым периодом, можно сделать вывод, что в выборке присутствуют большое количество либо совпадающих элементов, либо мало различающихся. Поскольку выборка является отрезком временного ряда, то именно соседние значения, как правило, практически не отличаются друг от друга. Этот факт негативно сказывается на дальнейшей кластеризации. Для того, чтобы уменьшить это отрицательное воздействие, для описания временного ряда мы будем использовать кусочно-постоянную

аппроксимацию. Суть аппроксимации заключается в следующем. Рассмотрим произвольный отрезок временного ряда $R_m^n = \{\rho_m, \rho_{m+1}, \dots, \rho_n\}$, который мы хотим заменить константным отрезком $\bar{R}_m^n = \{a_m^n, a_m^n, \dots, a_m^n\}$. Погрешность, которая при этом возникает, может быть измерена разными способами, например, при помощи $d_1 = \min_{a_m^n} \sqrt{\frac{1}{n+1-m} \sum_{i=m}^n (\rho_i - a_m^n)^2}$, или $d_2 = \min_{a_m^n} \frac{1}{n+1-m} \sum_{i=m}^n |\rho_i - a_m^n|$. Задача заключается в аппроксимации временного ряда наименьшим числом постоянных отрезков, достигнув при этом заданной точности. Эта задача относится к задачам, в которых решение может быть найдено в результате перебора всех вариантов. Приведенный ниже алгоритм находит приемлемое решение, но, возможно, неточное. Пусть задана точность ε .

Алгоритм.

1. Полагаем $m = n = 1$
2. Вычисляем d_1 или d_2
3. Если $d_1 \leq \varepsilon$ ($d_2 \leq \varepsilon$) то
 - $n = n + 1$
 - иначе
 - $m = n$
4. Если $n > N$ то Стоп
иначе перейти к 2.

Пример. Для расчетов были использованы данные цен акций компании Apple за две недели с периодом 17 секунд. Были вычислены относительные доходности ρ_i . Всего было использовано около 10000 записей о ценах акции. Данные были предварительно обработаны ($\varepsilon = 10^{-5}$) и разделены на два кластера, согласно приведенным выше параметрическому и непараметрическому методам, с вычислением доверительных интервалов с уровнем доверия $\alpha = 0.99$. В результате обработки 5000 записей были получены следующие кластеры: $[-0.00672, -0.00501]$ и $[0.00652, 0.00702]$, при этом предпочтение было отдано методу k -средних, поскольку гипотеза о нормальном распределении оказалась несостоятельной.

Рассмотрим приложение кластерного анализа к задачам финансовой математики.

Рассматривается (B, S) - рынок с рисковым активом S и безрисковым активом B . При стохастическом рассмотрении возврат рискового актива $\rho_n = \frac{\Delta S_n}{S_{n-1}}$ является случайной величиной, например, в модели КРР носитель распределения состоит из двух точек $\{a, b\}$. Для безрискового актива возврат $r = \frac{\Delta B_n}{B_{n-1}}$ является константой. Предполагается заданным стохастический базис: $\langle \Omega, (F_n), F \rangle$, на котором рассматривается множество мартингалльных мер

$\tilde{P} = \left\{ p : \left\langle \frac{S_n}{B_n}, F_n, p \right\rangle - \text{мартингал} \right\}$. Если множество мартингальных мер –

непустое множество, то величина $C^* = \frac{B_0}{B_N} \sup_{p \in \tilde{P}} E_p f$ называется верхней ценой

финансового обязательства f европейского опциона. Величина $C_* = \frac{B_0}{B_N} \inf_{p \in \tilde{P}} E_p f$

– нижняя цена финансового обязательства. Если существуют конечные значения для C_* и C^* , то $C_* \leq C^*$, и интервал $[C_*, C^*]$ называют интервалом справедливых цен. Поскольку цена $C < C_*$ создает арбитраж для покупателя, а цена $C > C^*$ создает арбитраж для продавца. Статистическая задача вычисления интервала справедливых цен имеет два существенных недостатка. Первый заключается в том, что длина интервала справедливых цен оказывается чрезмерно большой, второй заключается в том, что параметры модели требуется определить. Источником информации является выборка, с помощью которой определяются оценки параметров модели. Любая оценка является случайной величиной, следовательно, интервал справедливых цен является случайным.

Робастная постановка задачи вычисления оценки сверху справедливой цены европейского опциона является минимаксной задачей следующего вида:

$$\min_{\gamma} \max_{\rho} \left(\frac{f(S_N)}{B_N} - \frac{1}{1+r} \sum_{i=1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \right) \quad \text{при ограничениях:} \quad \rho_i \in M,$$

$$S_n = S_{n-1}(1 + \rho_n), B_n = B_{n-1}(1 + r). \text{ Предполагается, что начальные значения } S_0 \text{ и } B_0$$

– известны. Величина $C^* = B_0 \min_{\gamma} \max_{\rho} \left(\frac{f(S_N)}{B_N} - \frac{1}{1+r} \sum_{i=1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \right)$ является

верхней ценой европейского опциона с финансовым обязательством $f(S_N)$. Эта

цена обеспечивает достаточность капитала X_N^z самофинансируемого портфеля

для оплаты финансового обязательства при наихудшем развитии (B, S) - рынка,

то есть обеспечивает арбитраж для «продавца» опциона. В отличие от

«классического» определения верхней цены финансового обязательства, ρ не

является случайной величиной. Нижняя цена

$$C_* = B_0 \max_{\gamma} \min_{\rho} \left(\frac{f(S_N)}{B_N} - \frac{1}{1+r} \sum_{i=1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \right) \quad \text{при тех же ограничениях.}$$

Робастные методы представляют большой интерес для изучения. Робастные

методы, как правило, обладают устойчивостью по отношению к различного

рода помехам. Кроме этого, робастные методы снижают влияние случайных

выбросов на окончательный результат. Основу робастных методов составляют

робастные оптимизационные методы. Методы робастной выпуклой

оптимизации изложены в работах.

В силу большой актуальности, робастные методы имеют много

приложений. Одним из приложений является теория игр. Например, в

некоторых работах рассматриваются модели игр с неполной (конфиденциальной) информацией, выигрыш которых неопределён, и находится условие равновесия. Помимо теории игр, робастные модели имеют многочисленные приложения в финансовой математике. К примеру, иногда применяют робастную оптимизацию для оценки цен широкого спектра опционов. Инвесторы решают робастную оптимизационную задачу, решение которой определяет портфель, минимизирующий ошибку репликации финансового обязательства для наихудшего случая. Имеет смысл рассматривать задачу квантильного хеджирования финансовых обязательств, выплаты по которым неопределенны и находятся в некотором заданном диапазоне. Например, задачу робастной оптимизации для нахождения оптимального портфеля с инвестициями для модели с неопределёнными параметрами или робастную задачу нахождения оптимального портфеля в контексте торговли, требующей быстрого обновления портфеля при поступлении новых данных.

Рассмотрим робастную модель Кокса-Росса-Рубинштейна, в которой параметры классической модели a и b заменяются непересекающимися интервалами $[a_1, a_2]$ и $[a_3, a_4]$. Приведём решение задач вычисления верхней и нижней цен европейского и азиатского опционов, называемые задачами продавца и покупателя, для выпуклого финансового обязательства. Для вычисления интервалов применяется специальная технология обработки данных. Длина интервала справедливых цен в робастной модели зависит непосредственно от выборки и может регулироваться в зависимости от заданной надежности. Приведём результаты вычислений на реальных данных.

Базовые задачи. Здесь мы рассмотрим задачи, решения которых будут использованы далее в многошаговых моделях. Первая задача заключается в следующем:

$$\min \frac{x}{B_0}, \frac{x}{B_0} \geq \max_{\rho} \left[\frac{f(S_0(1+\rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1+r} \right], \rho \in [a_1, a_2] \cup [a_3, a_4], a_2 < a_3,$$

$$\Delta B_1 = B_0 r, a_2 < r < a_3.$$

Предполагается, что $f(y)$ – выпуклая функция. Так как для всех значений γ функция $\frac{f(S_0(1+\rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1+r}$ является выпуклой по переменной ρ , поэтому

$$\max_{\rho} \left[\frac{f(S_0(1+\rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1+r} \right] =$$

$$= \max \left[\frac{f(S_0(1+a_1))}{B_1} - \gamma \frac{S_0}{B_0} \frac{a_1 - r}{1+r}, \frac{f(S_0(1+a_4))}{B_1} - \gamma \frac{S_0}{B_0} \frac{a_4 - r}{1+r} \right]$$

В результате получаем задачу линейного программирования:

$$\min \frac{x}{B_0}, \frac{x}{B_0} + \gamma \frac{S_0}{B_0} \frac{a_i - r}{1+r} \geq \frac{f(S_0(1+a_i))}{B_1}, i = 1, 4.$$

Рассмотрим двойственную задачу линейного программирования. Поскольку в прямой задаче переменные x и γ могут быть как отрицательными, так и положительными, то в двойственной задаче ограничения – ограничения равенства: $y_1 + y_2 = 1, y_1 a_1 + y_2 a_4 = r$, поскольку в исходной задаче ограничения – неравенства (неважно какого знака), то двойственные переменные $y_i \geq 0$, наконец, целевая функция задачи $F = y_1 \frac{f(S_0(1+a_1))}{B_1} + y_2 \frac{f(S_0(1+a_4))}{B_1}$. Таким образом, двойственная задача заключается в следующем:

$$\max \left[y_1 \frac{f(S_0(1+a_1))}{B_1} + y_2 \frac{f(S_0(1+a_4))}{B_1} \right], y_1 + y_2 = 1, y_1 a_1 + y_2 a_4 = r, y_i \geq 0.$$

Система линейных уравнений имеет единственное решение:

$$y_1 = \frac{a_4 - r}{a_4 - a_1}, y_2 = \frac{r - a_1}{a_4 - a_1}. \text{ Это решение будет допустимым так как } a_1 < r < a_4.$$

Таким образом, оптимальное значение

$$x^* = \frac{1}{1+r} \left[\frac{a_4 - r}{a_4 - a_1} f(S_0(1+a_1)) + \frac{r - a_1}{a_4 - a_1} f(S_0(1+a_4)) \right].$$

Из условий Куна-Таккера вычисляется оптимальное значение для

$$\gamma^* = \frac{f(S_0(1+a_4)) - f(S_0(1+a_1))}{S_0(a_4 - a_1)}.$$

Рассмотрим вторую задачу:

$$\max \frac{x}{B_0}, \frac{x}{B_0} \leq \min_{\rho} \left[\frac{f(S_0(1+\rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1+r} \right]; \rho \in [a_1, a_2] \cup [a_3, a_4], a_2 < a_3;$$

$$\Delta B_1 = B_0 r; a_2 < r < a_3.$$

Рассмотрим дискретную аппроксимацию задачи:

$$\frac{x}{B_0} + \gamma \frac{S_0}{B_0} \frac{\rho_i - r}{1+r} \leq \frac{f(S_0(1+\rho_i))}{B_1}, \max \frac{x}{B_0};$$

которая возникает в результате кусочно-постоянной аппроксимации непрерывной на компакте функции. Пусть $\rho_{i^*} = a_2, \rho_{i^*+1} = a_3$. Рассмотрим двойственную задачу:

$$\sum_i y_i = 1, \sum_i \rho_i y_i = r, y_i \geq 0, \min F = \frac{1}{B_1} \sum_i y_i f(S_0(1+\rho_i)).$$

Выберем в качестве базисных переменных y_{i^*}, y_{i^*+1} . Соответствующее

базисное решение $y_{i^*} = \frac{a_3 - r}{a_3 - a_2}, y_{i^*+1} = \frac{r - a_2}{a_3 - a_2}, y_i = 0, i \neq i^*, i \neq i^* + 1$ является допустимым.

Выразим целевую функцию через свободные переменные:

$$B_1 \cdot F = \frac{a_3 - r}{a_3 - a_2} f(S_0(1 + a_2)) + \frac{r - a_2}{a_3 - a_2} f(S_0(1 + a_3)) - \sum_{i \neq i^*, i \neq i^* + 1} \left[\frac{a_3 - \rho_i}{a_3 - a_2} f(S_0(1 + a_2)) + \frac{\rho_i - a_2}{a_3 - a_2} f(S_0(1 + a_3)) - f(S_0(1 + \rho_i)) \right] y_i.$$

Пусть $\rho_i > a_3$. Выразим $a_3 = \alpha a_2 + (1 - \alpha)\rho_i$, отсюда $\alpha = \frac{\rho_i - a_3}{\rho_i - a_2}$.

Поскольку f – выпуклая функция, то справедливо неравенство: $f(S_0(1 + a_3)) \leq \alpha f(S_0(1 + a_2)) + (1 - \alpha)f(S_0(1 + \rho_i))$. Отсюда

$\frac{1}{1 - \alpha} f(S_0(1 + a_3)) - \frac{\alpha}{1 - \alpha} f(S_0(1 + a_2)) \leq f(S_0(1 + \rho_i))$. После подстановки α в данное неравенство получим

$$\frac{a_3 - \rho_i}{a_3 - a_2} f(S_0(1 + a_2)) + \frac{\rho_i - a_2}{a_3 - a_2} f(S_0(1 + a_3)) \leq f(S_0(1 + \rho_i)). \quad \text{Аналогично}$$

устанавливается, что рассматриваемое неравенство справедливо и для случая $\rho_i < a_2$. Следовательно, коэффициенты при свободных переменных отрицательны. Поэтому рассматриваемое базисное решение является оптимальным. Отсюда

$$x^* = \frac{1}{1 + r} \left[\frac{a_3 - r}{a_3 - a_2} f(S_0(1 + a_2)) + \frac{r - a_2}{a_3 - a_2} f(S_0(1 + a_3)) \right],$$

$$\gamma^* = \frac{f(S_0(1 + a_3)) - f(S_0(1 + a_2))}{S_0(a_3 - a_2)}.$$

Заметим, что решение не зависит ни от точек разбиения, ни от их числа. Устремив число точек разбиения к бесконечности таким образом, чтобы диаметр разбиения стремился к нулю, мы получим решение непрерывной задачи.

Третья и четвертая задачи заключается в следующем. Требуется найти

$$\min \frac{x}{B_0}, \frac{x}{B_0} \geq \max_{\rho} \left[\frac{f(S_0 + S_0(1 + \rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1 + r} \right], \rho \in [a_1, a_2] \cup [a_3, a_4], a_2 < a_3$$

$$\Delta B_1 = B_0 r, a_2 < r < a_3.$$

$$\max \frac{x}{B_0}, \frac{x}{B_0} \leq \min_{\rho} \left[\frac{f(S_0 + S_0(1 + \rho))}{B_1} - \gamma \frac{S_0}{B_0} \frac{\rho - r}{1 + r} \right]; \rho \in [a_1, a_2] \cup [a_3, a_4], a_2 < a_3;$$

$$\Delta B_1 = B_0 r; a_2 < r < a_3.$$

Третья и четвертая задачи эквивалентны первой и второй задачам, поэтому мы приведем их решение без необходимых расчетов:

$$x^* = \frac{1}{1+r} \left[\frac{a_4 - r}{a_4 - a_1} f(S_0 + S_0(1+a_1)) + \frac{r - a_1}{a_4 - a_1} f(S_0 + S_0(1+a_4)) \right],$$

$$\gamma^* = \frac{f(S_0 + S_0(1+a_4)) - f(S_0 + S_0(1+a_1))}{S_0(a_4 - a_1)} \quad - \text{ решение третьей задачи;}$$

$$x^* = \frac{1}{1+r} \left[\frac{a_3 - r}{a_3 - a_2} f(S_0 + S_0(1+a_2)) + \frac{r - a_2}{a_3 - a_2} f(S_0 + S_0(1+a_3)) \right],$$

$$\gamma^* = \frac{f(S_0 + S_0(1+a_3)) - f(S_0 + S_0(1+a_2))}{S_0(a_3 - a_2)} \quad - \text{ решение четвертой задачи.}$$

Многошаговые задачи. Рассмотрим многошаговый вариант первой задачи:

$$X_0(x) = \min_{\gamma^N} y, \frac{y}{B_0} + \frac{1}{1+r} \sum_{i=1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \geq \frac{f\left(x \prod_{i=1}^N (1 + \rho_i)\right)}{B_N}, \forall \rho_i \in [a_1, a_2] \cup [a_3, a_4],$$

$\Delta S_i = S_{i-1} \rho_i, \Delta B_i = B_{i-1} r, S_0 = x$. Здесь использовано обозначение $\gamma_r^s = \{\gamma_r, \gamma_{r+1}, \dots, \gamma_s\}$. Многошаговая задача является задачей динамического программирования. Для ее решения рассмотрим в момент времени k функцию Беллмана $X_k(x)$:

$$X_k(x) = \min_{\gamma_{k+1}} y,$$

$$\frac{y}{B_k} + \frac{1}{1+r} \sum_{i=k+1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \geq \frac{f\left(x \prod_{i=k+1}^N (1 + \rho_i)\right)}{B_N}, \forall \rho_i \in [a_1, a_2] \cup [a_3, a_4], S_i = S_{i-1}(1 + \rho_i),$$

$$B_i = B_{i-1}(1+r), S_k = x.$$

То есть для того, чтобы вычислить значение функции $X_k(x)$, которую далее будем называть капиталом, требуется решить задачу оптимизации с бесконечным числом ограничений. Тем не менее, вычисление значения $X_k(x)$ определяется несложной формулой, которая является результатом следующей теоремы.

Теорема (Задача продавца). Если $f(x)$ – выпуклая функция, то для капитала $X_k(x)$ справедлива следующая формула:

$$X_k(x) = \frac{1}{(1+r)^{N-k}} \sum_{j=0}^{N-k} C_{N-k}^j f(x(1+a_4)^j (1+a_1)^{N-k-j}) (p^*)^j (q^*)^{N-k-j},$$

$$p^* = \frac{r - a_1}{a_4 - a_1}, q^* = 1 - p^*.$$

Теорему несложно доказать методом математической индукции, используя результаты, полученные для одношаговой модели. А именно, формулу

$$X_k(x) = \frac{1}{1+r} (X_{k+1}(x(1+a_4))p^* + X_{k+1}(x(1+a_1))q^*).$$

Очевидно, что для вычисления портфеля следует использовать формулу:

$$\gamma_{k+1}^*(x) = \frac{X_{k+1}(x(1+a_4)) - X_{k+1}(x(1+a_1))}{x(a_4 - a_1)}.$$

Аналогичная теорема справедлива для задачи покупателя:

$$X_0(x) = \max_{\gamma_i^N} y, \frac{y}{B_0} + \frac{1}{1+r} \sum_{i=1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \leq \frac{f\left(x \prod_{i=1}^N (1+\rho_i)\right)}{B_N}, \forall \rho_i \in [a_1, a_2] \cup [a_3, a_4],$$

$$\Delta S_i = S_{i-1} \rho_i, \Delta B_i = B_{i-1} r, S_0 = x.$$

Определим капитал $X_k(x)$ для задачи покупателя:

$$X_k(x) = \max_{\gamma_{k+1}^N} y, \frac{y}{B_k} + \frac{1}{1+r} \sum_{i=k+1}^N \gamma_i \frac{S_{i-1}}{B_{i-1}} (\rho_i - r) \leq \frac{f\left(x \prod_{i=k+1}^N (1+\rho_i)\right)}{B_N}, \forall \rho_i \in [a_1, a_2] \cup [a_3, a_4],$$

$$\Delta S_i = S_{i-1} \rho_i, B_i = B_{i-1} r, S_k = x.$$

Теорема (Задача покупателя). Если $f(x)$ – выпуклая функция, то для капитала $X_k(x)$ справедлива следующая формула:

$$X_k(x) = \frac{1}{(1+r)^{N-k}} \sum_{j=0}^{N-k} C_{N-k}^j f(x(1+a_2)^j (1+a_3)^{N-k-j}) (p^*)^j (q^*)^{N-k-j},$$

$$p^* = \frac{r - a_2}{a_3 - a_2}, q^* = 1 - p^*.$$

Для вычисления портфеля следует использовать формулу:

$$\gamma_{k+1}^*(x) = \frac{X_{k+1}(x(1+a_3)) - X_{k+1}(x(1+a_2))}{x(a_3 - a_2)}.$$

Рассмотрим азиатский опцион с платёжной функцией

$$f_N = \left(\frac{1}{N+1} \sum_{i=0}^N S_i - K \right)^+.$$

Данную платёжную функцию можно рассматривать как

частный случай более общей функции $f_N = f\left(\sum_{i=0}^N S_i\right)$, где $f(x)$ – выпуклая функция. То есть в случае азиатского опциона значение платёжной функции зависит от всех значений стоимости акции. В случае европейского опциона значение платёжной функции зависит только от последнего значения стоимости акции.

Рекуррентные формулы для расчёта оптимального капитала и портфеля в случае задачи продавца:

$$X_k(u, v) = \frac{1}{1+r} [q^* X_{k+1}(u(1+a_1), v + u(1+a_1)) + p^* X_{k+1}(u(1+a_4), v + u(1+a_4))],$$

$$X_N(v) = f(v) = \left(\frac{v}{N+1} - K \right)^+, p^* = \frac{r - a_1}{a_4 - a_1}, q^* = 1 - p^*,$$

$$\gamma_{k+1}(u, v) = \frac{X_{k+1}(u(1+a_4), v + u(1+a_4)) - X_{k+1}(u(1+a_1), v + u(1+a_1))}{u(a_4 - a_1)}.$$

Аналогичные формулы справедливы для задачи покупателя.

Лекция 5.

«Линейные стохастические модели. Модель скользящего среднего МА,
авторегрессионная модель AR»

Рассмотрим модель $MA(q)$: $h_n = \mu + \beta(L)\varepsilon_n, \beta(L) = b_0 + b_1L + \dots + b_qL^q$.

Легко находим, что $Eh_n = \mu, Dh_n = b_0^2 + b_1^2 + \dots + b_q^2$ и $R(k) = \begin{cases} \sum_{j=0}^{q-k} b_j b_{k+j}, & k = 1, \dots, q \\ 0, & k > q \end{cases}$.

Видно, что схемами типа $MA(q)$ можно путём варьирования коэффициентов b_i пытаться моделировать поведение последовательностей $h = (h_n)$, у которых корреляция величин $h_n, h_{n+k}, k > q$ и равна нулю. Причём под моделированием поведения последовательности мы будем понимать подгонку той или иной модели к эмпирическим данным. Опишем общий алгоритм.

По выборочным значениям h_1, h_2, \dots находим выборочное среднее $\bar{h}_n = \frac{1}{n} \sum_{k=1}^n h_k$, выборочную дисперсию $\bar{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^n (h_k - \bar{h}_n)^2$, выборочную

корреляцию порядка k : $r_n(k) = \frac{1}{n\bar{\sigma}_n^2} \sum_{i=k+1}^n (h_i - \bar{h}_n)(h_{i-k} - \bar{h}_n)$, выборочные частичные

корреляции и т.д. Затем, используя выражения для соответствующих теоретических характеристик аппроксимируемых моделей, производится варьирование параметров с целью подгонки теоретических характеристик под эмпирические. Наконец, на заключительном этапе производится оценка качества подгонки, основываясь на знании распределений эмпирических характеристик и их отклонений от теоретических распределений.

Рассмотрим модель $MA(\infty)$, в которой $h_n = \mu + \sum_{j=0}^{\infty} b_j \varepsilon_{n-j}$. Если $\sum_{j=0}^{\infty} b_j^2 < \infty$, то ряд будет сходиться в среднеквадратическом смысле. Имеем:

$$Eh_n = \mu, Dh_n = \sum_{j=0}^{\infty} b_j^2, R(k) = \sum_{j=0}^{\infty} b_{k+j} b_j, k \geq 0.$$

В теории стационарных случайных процессов принято говорить, что h_n есть результат реакции физически осуществимого фильтра с импульсной переходной функцией $b = (b_j)$, когда на вход подаётся последовательность $\varepsilon = (\varepsilon_n)$.

Модель $AR(p)$ имеет вид: $h_n = a_0 + a_1 h_{n-1} + \dots + a_p h_{n-p} + \sigma \varepsilon_n$. Если $w_n = a_0 + \sigma \varepsilon_n$, то $w_n = (1 - a_1 L - a_2 L^2 - \dots - a_p L^p) h_n$. Предположим, что найдено факторизационное представление:

$$1 - a_1 L - a_2 L^2 - \dots - a_p L^p = (1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L)$$

с разными $\lambda_i, i = 1, \dots, p$. Если

$|\lambda_i| < 1, i = 1, \dots, p$, то стационарное решение, единственное в классе решений с конечным вторым моментом, идущих по времени из $-\infty$: $h_n = (1 - \lambda_1 L)^{-1} \dots (1 - \lambda_p L)^{-1} w_n$. Числа $\lambda_1, \dots, \lambda_p$ являются корнями уравнения:

$\lambda^p - a_1\lambda^{p-1} - \dots - a_{p-1}\lambda - a_p = 0$. Равносильным образом, можно сказать, что $\lambda_i = z_i^{-1}$, где z_i – корни уравнения $1 - a_1z - a_2z^2 - \dots - a_pz^p = 0$. Стационарное решение мы получаем, если все корни этого уравнения лежат вне единичного круга.

Пусть $\frac{1}{(1-\lambda_1z)\dots(1-\lambda_pz)} = \frac{c_1}{1-\lambda_1z} + \dots + \frac{c_p}{1-\lambda_pz}$. Найдём константы c_1, \dots, c_p .

Умножив обе части равенства на $(1-\lambda_1z)\dots(1-\lambda_pz)$, получим, что $1 = \sum_{i=1}^p c_i \prod_{1 \leq k \leq p, k \neq i} (1-\lambda_kz)$. Это равенство должно быть выполнено для всех z . Значит,

и для $z = \lambda_1^{-1}, \dots, z = \lambda_p^{-1}$. Тогда $c_i = \frac{\lambda_i^{p-1}}{\prod_{\substack{1 \leq k \leq p, \\ k \neq i}} (\lambda_i - \lambda_k)}$. Заметим, что $\sum_{i=1}^p c_i = 1$. Имеем:

$h_n = \sum_{l=0}^{\infty} (c_1\lambda_1^l + \dots + c_p\lambda_p^l) w_{n-l}$. Данное представление даёт возможность подсчитывать различные характеристики последовательности $h = (h_n)$. Например, моменты Eh_n^k , ковариации, условные математические ожидания $E(h_{n+k} / F_n)$, где $F_n = \sigma(\dots, w_{-1}, w_0, \dots, w_n)$ и т.д.

В предположении стационарности моменты $Eh_n \equiv \mu$ имеют вид:

$\mu = \frac{a_0}{1 - (a_1 + \dots + a_p)}$. Ковариации $R(k) = \text{cov}(h_n, h_{n+k})$ для $k = 1, 2, \dots$ имеют вид:

$R(k) = a_1R(k-1) + \dots + a_pR(k-p)$. Если $k = 0$, то $R(0) = a_1R(1) + \dots + a_pR(p) + \sigma^2$. Для корреляционных функций $\rho(k), k \geq 0$ справедливы аналогичные уравнения, называемые уравнениями Юла-Уолкера.

Дополнительную информацию о свойствах оценок максимального правдоподобия можно получить из следующих результатов. Во-первых, $\sup_{a_1 \in R} E_{a_1} |\hat{a}_1 - a_1| \rightarrow 0, n \rightarrow \infty$. Пусть $U(a_1)$ – класс оценок \tilde{a}_1 , для которых

смещение $b_{a_1}(\tilde{a}_1) \equiv E_{a_1}(\tilde{a}_1 - a_1) \rightarrow 0, n \rightarrow \infty$ и $\frac{db_{a_1}}{da_1}(\tilde{a}_1) \rightarrow 0, n \rightarrow \infty$. Оценки

максимального правдоподобия \hat{a}_1 принадлежат классу $U(a_1)$ при $|a_1| \neq 1$.

Для $|a_1| \neq 1$ оценки максимального правдоподобия \hat{a}_1 являются асимптотически эффективными в классе $U(a_1)$ в том смысле, что для всех

$\tilde{a}_1 \in U(a_1)$ $\overline{\lim}_n \frac{E_{a_1} \langle M_n \rangle (\hat{a}_1 - a_1)^2}{E_{a_1} \langle M_n \rangle (\tilde{a}_1 - a_1)^2} \leq 1$. В стационарном случае (когда $|a_1| < 1$)

оценки \hat{a}_1 асимптотически эффективны также и в обычном смысле: для всех

$\tilde{a}_1 \in U(a_1)$ $\overline{\lim}_n \frac{D_{a_1} \hat{a}_1}{D_{a_1} \tilde{a}_1} \leq 1$.

Последовательные оценки максимального правдоподобия обладают следующими свойствами асимптотической равномерности: при $\theta \rightarrow \infty$

$$\sup_{|a_1| \leq 1} \sup_x \left| P_{a_1} \left\{ \sqrt{\langle M \rangle_{r(\theta)}} [\tilde{a}(r(\theta)) - a_1] \leq x \right\} - \Phi(x) \right| \rightarrow 0$$

$$\sup_{1 < r \leq |a_1| \leq R} \sup_x \left| P_{a_1} \left\{ \sqrt{\langle M \rangle_{r(\theta)}} [\hat{a}(r(\theta)) - a_1] \leq x \right\} - \Phi(x) \right| \rightarrow 0.$$

Один из центральных вопросов статистики авторегрессионных схем $AR(p)$ состоит в оценивании параметров $\theta = (a_0, a_1, \dots, a_p, \sigma)$. Для определённости будем считать h_0, h_{-1}, \dots известными константами.

Если предполагать, что белый шум в широком смысле $\varepsilon = (\varepsilon_n)$ является гауссовским, то основным методом оценивания является метод максимального правдоподобия, согласно которому в качестве оценки параметра θ по наблюдениям h_1, h_2, \dots, h_n берётся значение $\hat{\theta}_n = \arg \max_{\theta} p_{\theta}(h_1, h_2, \dots, h_n)$, где $p_{\theta}(h_1, h_2, \dots, h_n)$ – совместная плотность гауссовского вектора (h_1, h_2, \dots, h_n) .

Общие принципы метода максимального правдоподобия проиллюстрируем на примере модели $AR(1)$, в которой $h_n = a_0 + a_1 h_{n-1} + \sigma \varepsilon_n$, причём $\sigma > 0$ будем считать известным параметром, $h_0 = 0, n \geq 1$.

Поскольку здесь $\theta = (a_0, a_1)$, то

$$p_{\theta}(h_1, \dots, h_n) = \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^n \exp \left(-\frac{1}{2} \sum_{k=1}^n \frac{(h_k - a_0 - a_1 h_{k-1})^2}{2\sigma^2} \right). \quad \text{Оценка} \quad \hat{\theta} = (\hat{a}_0, \hat{a}_1)$$

определяется из условия обращения в минимум функции

$\psi(a_0, a_1) = \sum_{k=1}^n (h_k - a_0 - a_1 h_{k-1})^2$. Считая, что $a_0 \in R, a_1 \in R$, находим, что

$$\frac{\partial \psi}{\partial a_0} = 0 \Leftrightarrow 2 \sum_{k=1}^n (h_k - a_0 - a_1 h_{k-1}) = 0$$

$$\frac{\partial \psi}{\partial a_1} = 0 \Leftrightarrow 2 \sum_{k=1}^n (h_k - a_0 - a_1 h_{k-1}) h_{k-1} = 0$$

Решая эту линейную систему, находим оценки \hat{a}_0, \hat{a}_1 .

Рассмотрим свойства оценок $\hat{a}_1 = \hat{a}_1(h_1, \dots, h_n), n \geq 1$ параметра a_1 , считая параметр $a_0 = 0, \sigma = 1$.

$$\text{В этом предположении } h_n = a_1 h_{n-1} + \varepsilon_n, \text{ и } \hat{a}_1 = \frac{\sum_{k=1}^n h_{k-1} h_k}{\sum_{k=1}^n h_{k-1}^2} = a_1 + \frac{\sum_{k=1}^n h_{k-1} \varepsilon_k}{\sum_{k=1}^n h_{k-1}^2}.$$

Обозначим $M_n = \sum_{k=1}^n h_{k-1} \varepsilon_k$. Последовательность $M = (M_n)$ является при любом значении параметра a_1 мартингалом, квадратическая характеристика которого $\langle M \rangle_n = \sum_{k=1}^n h_{k-1}^2$. Поэтому, если a_1 – истинное значение неизвестного

параметра, то $\hat{a}_1 = a_1 + \frac{M_n}{\langle M \rangle_n}$. Из этого представления мы видим, что оценка максимального правдоподобия \hat{a}_1 является сильно состоятельной, то есть с вероятностью единица $\hat{a}_1 \rightarrow a_1, n \rightarrow \infty$, поскольку $\langle M \rangle_n \rightarrow \infty (P - \text{п.н.})$ и, согласно усиленному закону больших чисел для квадратично интегрируемых мартингалов $\frac{M_n}{\langle M \rangle_n} \rightarrow 0 (P - \text{п.н.})$.

Если при $a_0 = 0, \sigma = 1$ подсчитать информацию Фишера $I_n(a_1) = E_{a_1} \left\{ -\frac{\partial^2 \ln P_{a_1}(h_1, \dots, h_n)}{\partial a_1^2} \right\}$, где E_{a_1} – математическое ожидание по мере

$P_{a_1} = \text{Law}(h_1, \dots, h_n / \theta = a_1)$, то найдём, что $I_n(a_1) = E_{a_1} \langle M \rangle_n = E_{a_1} \sum_{k=1}^n h_{k-1}^2$. Используя

формулы для Eh_{k-1}, Dh_{k-1} , получим, что $I_n(a_1) \sim \begin{cases} \frac{n}{1-a_1^2}, |a_1| < 1 \\ \frac{n^2}{2}, |a_1| = 1 \\ \frac{a_1^{2n}}{(a_1^2-1)^2}, |a_1| > 1 \end{cases}$. В случае $|a_1| = 1$

$h = (h_n)$ является случайным блужданием. При $|a_1| \neq 1$ $h = (h_n)$ является марковской. При $|a_1| < 1, n \rightarrow \infty$ $h = (h_n)$ стационаризуется.

В статистике последовательностей $h = (h_n)$ большое внимание уделяется вопросу о том, какая из двух гипотез: $H_0: |a_1| = 1, H_1: |a_1| > 1$ является более правдоподобной. При этом вопрос, будет ли $|a_1| = 1$, носит название «проблемы единичного корня».

Приведём некоторые результаты относительно предельного распределения $\hat{a}_1 - a_1$.

Теорема 1. При $n \rightarrow \infty$ $\lim_n P_{a_1} \{ \sqrt{\langle M \rangle_n}^{-1} \sum_{k=1}^n h_{k-1} \leq x \} = \begin{cases} \Phi(x), |a_1| < 1 \\ H_{a_1}(x), |a_1| = 1, \\ Ch(x), |a_1| > 1 \end{cases}$ где

$\Phi(x) = \int_{-\infty}^x \varphi_{(0,1)}(y) dy$ – стандартное нормальное распределение, $Ch(x)$ – распределение Коши с плотностью $\frac{1}{\pi(1+x^2)}$ и $H_{a_1}(x)$ – распределение

случайной величины $a_1 \frac{W^2(1) - 1}{2\sqrt{2} \int_0^1 W^2(s) ds}$, где $(W_s)_{s \leq 1}$ – стандартный винеровский

процесс (броуновское движение).

Отметим, что при $|a_1| \neq 1$ плотности предельных распределений являются симметричными относительно нуля. Однако, плотность распределения $H_{a_1}(x)$ несимметрична. Это следует из того наблюдения, что $P(W^2(1) - 1 > 0) \neq \frac{1}{2}$.

Следующий результат показывает, что с помощью случайной нормировки отклонения $(\hat{a}_1 - a_1)$, состоящей в том, что вместо информации Фишера $I_n(a_1)$ используется стохастическая фишеровская информация $\langle M \rangle_n$, можно получить лишь только два предельных распределения вместо трёх.

Теорема 2. При $n \rightarrow \infty$ $\lim P_{a_1} \left\{ \sqrt{\langle M \rangle_n} (\hat{a}_1 - a_1) \leq x \right\} = \begin{cases} \Phi(x), & |a_1| \neq 1 \\ H'_{a_1}(x), & |a_1| = 1 \end{cases}$, где $H'_{a_1}(x)$ – распределение вероятностей случайной величины $a_1 \frac{W^2(1) - 1}{2 \sqrt{\int_0^1 W^2(s) ds}}$.

Использование наряду со случайной нормировкой последовательных оценок максимального правдоподобия приводит к единственному предельному (нормальному) распределению.

Теорема 3. Пусть $\theta > 0$, $\tau(\theta) = \inf \{n \geq 1 : \langle M \rangle_n \geq \theta\}$ и $\hat{a}_1(r(\theta)) = \frac{\sum_{k=1}^{\tau(\theta)} h_{k-1} h_k}{\sum_{k=1}^{\tau(\theta)} h_{k-1}^2}$ –

последовательная оценка максимального правдоподобия. Тогда при любом значении $a_1 \in R$ $\lim_{n \rightarrow \infty} P_{a_1} \left\{ \sqrt{\langle M \rangle_{r(\theta)}} |\hat{a}_1(r(\theta)) - a_1| \leq x \right\} = \Phi(x)$.

Если исходным временным параметром является время n , то параметр θ можно рассматривать как новое, операционное время, построенное по стохастической фишеровской информации $\langle M \rangle$. Заметим, что тем большим временным интервалам, где фишеровская информация мало меняется, в новом θ -времени будут соответствовать малые интервалы, и наоборот. Тем самым, в новом, операционном времени приток информации становится равномерным, однородным. В новом θ -времени поступающие данные становятся равноценными, одинаково распределёнными для всех значений параметра a_1 , что и определило то, что предельное распределение оказалось единственным и нормальным. Переход к новому времени производится с целью выравнивания статистических данных обменных курсов валют, имеющих в своём поведении географические компоненты периодического характера.

Рассмотрим вопрос прогнозирования в линейных моделях.

В линейных стационарных моделях имеется теория построения оптимальных (в среднеквадратическом смысле) линейных оценок, развитая в работах А.Н.Колмогорова и Н.Винера.

Многие из рассмотренных последовательностей $h = (h_n)$ могут быть представлены в виде одностороннего скользящего среднего $h_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$ с $\sum_{k=0}^{\infty} |a_k|^2 < \infty$ и некоторой базовой последовательностью $\varepsilon = (\varepsilon_n)$, являющейся белым шумом в широком смысле.

Приведём необходимые определения и понятия.

Если $\xi = (\xi_n)$ – стохастическая последовательность, то будем обозначать $F_n^\xi = \sigma(\dots, \xi_{n-1}, \xi_n)$ σ -алгебру, порождённую прошлым $\{\xi_k, k \leq n\}$, $F_{\infty-}^\xi = \vee F_n^\xi$ σ -алгебру, порождённую всеми величинами $\xi = (\xi_n)$, H_n^ξ – замкнутое (в L^2) линейное многообразие $\bar{L}(\dots, \xi_{n-1}, \xi_n)$, порождённое величинами $\{\xi_k, k \leq n\}$, H_∞^ξ – замкнутое линейное многообразие, порождённое всеми величинами $\xi = (\xi_n)$, то есть $H_\infty^\xi = \bar{L}(\xi_k, k < \infty)$.

Пусть $\eta = \eta(\omega)$ – некоторая случайная величина с конечным вторым моментом, $E\eta^2(\omega) < \infty$. Сформулируем задачу оценивания величины η по наблюдениям за последовательностью ξ .

Здесь наиболее распространены следующие два подхода.

Если наблюдению подлежат значения $(\dots, \xi_{n-1}, \xi_n)$, то в первом подходе в качестве оценок берётся класс всех F_n^ξ -измеримых оценок $\tilde{\eta}_n$ и среди них наилучшей (оптимальной) объявляется та оценка $\hat{\eta}_n$, на которой достигается минимум среднеквадратического отклонения: $E|\eta - \hat{\eta}_n|^2 = \inf_{\tilde{\eta}_n} E|\eta - \tilde{\eta}_n|^2$. Хорошо известно, что оптимальная в этом смысле оценка имеет следующий вид: $\hat{\eta}_n = E(\eta / F_n^\xi)$, где $E(\cdot, \cdot)$ – условное математическое ожидание, являющееся нелинейной функцией от наблюдаемых значений $(\dots, \xi_{n-1}, \xi_n)$.

В другом подходе к рассмотрению допускаются лишь линейные оценки, то есть функции из H_n^ξ , образованные или конечными линейными комбинациями величин $\{\xi_k, k \leq n\}$, или их L^2 -замыканиями.

Наилучшей (оптимальной) линейной оценкой объявляется оценка $\hat{\lambda}_n \in H_n^\xi$ такая, что $E|\eta - \hat{\lambda}_n|^2 = \inf_{\tilde{\lambda}_n \in H_n^\xi} E|\eta - \tilde{\lambda}_n|^2$. При этом для $\hat{\lambda}_n$ принято следующее обозначение: $\hat{\lambda}_n = \hat{E}(\eta / H_n^\xi)$, где $\hat{E}(\cdot, \cdot)$ – условное математическое ожидание в широком смысле.

Вернёмся к вопросу о (линейном) предсказании значения $h_n, n \geq 1$, и последовательности $h = (h_n)$ по той информации, которая имеется до момента времени 0. Сравнительно просто эта задача решается, если информацией считается все прошлое, порождённое не $h = (h_n)_{n \leq 0}$, а $\varepsilon = (\varepsilon_n)_{n \leq 0}$: $H_0^\varepsilon = \bar{L}(\dots, \varepsilon_{-1}, \varepsilon_0)$. Имеем:

$$\hat{h}_n^{(\varepsilon)} = \hat{E}(h_n / H_0^\varepsilon) = \hat{E}\left(\sum_{k=0}^{n-1} a_k \varepsilon_{n-k} + \sum_{k=n}^{\infty} a_k \varepsilon_{n-k} / H_0^\varepsilon\right) = \sum_{k=n}^{\infty} a_k \varepsilon_{n-k} + \hat{E}\left(\sum_{k=0}^{n-1} a_k \varepsilon_{n-k} / H_0^\varepsilon\right) =$$

$= \sum_{k=n}^{\infty} a_k \varepsilon_{n-k}$, так как для $i \geq 1$ $\hat{E}(\varepsilon_i / H_0^\varepsilon) = E\varepsilon_i = 0$, что следует из ортогональности компонент в белом шуме $\varepsilon = (\varepsilon_i)$.

При этом ошибка экстраполяции: $\sigma_n^2 = E|h_n - \hat{E}(h_n / H_0^\varepsilon)|^2 = \sum_{k=0}^{n-1} |a_k|^2$. Ясно, что $\sigma_n^2 \leq \sigma_{n+1}^2$ и $\lim_n \sigma_n^2 = \sum_{k=0}^{\infty} |a_k|^2 = Eh_n^2$. Отсюда следует, что с ростом n роль прошлой информации $H_0^\varepsilon = \bar{L}(\dots, \varepsilon_{n-1}, \varepsilon_0)$ в предсказании значений величины h_n становится все меньше и меньше, и в пределе при $n \rightarrow \infty$ в качестве наилучшей линейной оценки надо брать просто математическое ожидание, то есть 0 в рассматриваемом случае.

Большой интерес представляет задача экстраполяции не только по бесконечному числу значений $\{\varepsilon_k, k \leq 0\}$, но и по некоторому их конечному числу $\{\varepsilon_k, l \leq k \leq 0\}$. Представление $\hat{h}_n^{(\varepsilon)} = \sum_{k=n}^{\infty} a_k \varepsilon_{n-k}$ решает задачу экстраполяции величин h_n по линейной информации H_0^ε , а не по информации H_0^h , содержащейся в $\bar{L}(\dots, h_{-1}, h_0)$. Если предположить, что $H_0^\varepsilon = H_0^h$, то можно получить оценку $\hat{E}(h_n / H_0^h)$ в виде $\sum_{k=n}^{\infty} \hat{a}_k h_{n-k}$.

Напомним некоторые общие результаты теории стационарных случайных последовательностей.

Пусть $\xi = (\xi_n)$ – некоторая стационарная последовательность.

Каждый элемент $\eta \in H(\xi)$ может быть представлен в виде суммы двух ортогональных компонент: $\eta = \hat{E}(\eta / S(\xi)) + [\eta - \hat{E}(\eta / S(\xi))]$, где $S(\xi) = \bigcap_n H_n^\xi$ – линейная информация, содержащаяся в бесконечно далёком прошлом. Обозначая $R(\xi)$ множество элементов вида $\eta - \hat{E}(\eta / S(\xi))$, где $\eta \in H(\xi)$, находим, что и само множество $H(\xi)$ может быть представлено как ортогональная сумма: $H(\xi) = S(\xi) \oplus R(\xi)$.

Стационарная последовательность $\xi = (\xi_n)$ называется регулярной, если $H(\xi) = R(\xi)$, и сингулярной, если $H(\xi) = S(\xi)$. Условие $H(\xi) = S(\xi)$ означает, что вся информация, доставляемая значениями последовательности ξ , находится в бесконечно далёком прошлом. Поэтому сингулярные последовательности называют также чисто (вполне) детерминированными. Если подпространство $S(\xi) = \emptyset$, последовательность ξ называют чисто (вполне) недетерминированной.

Предложение 1. Всякая невырожденная стационарная в широком смысле случайная последовательность $\xi = (\xi_n)$ допускает, и притом единственное, разложение $\xi_n = \xi_n^r + \xi_n^s$, где $\xi^r = (\xi_n^r)$ – регулярная последовательность, $\xi^s = (\xi_n^s)$ – сингулярная последовательность. При этом ξ^r и ξ^s ортогональны, то есть для всех n, m $\text{cov}(h_n^r, h_m^s) = 0$.

Определение. Случайная последовательность $\varepsilon = (\varepsilon_n)$ называется обновляющей последовательностью для $\xi = (\xi_n)$, если:

1. $\varepsilon = (\varepsilon_n)$ является белым шумом в широком смысле, то есть $E\varepsilon_n = 0, E\varepsilon_n \varepsilon_m = 0, m \neq n, E|\varepsilon_n|^2 = 1$.
2. Для всех n $H_n^\varepsilon = H_n^\xi$.

В силу ортогональности элементов последовательности $\varepsilon = (\varepsilon_n)$ величина ε_{n+1} может рассматриваться как та новая информация, которая обновляет информацию, содержащуюся в H_n^ε , и даёт возможность образовать H_{n+1}^ε . Поскольку для всех n $H_n^\varepsilon = H_n^\xi$, то, тем самым, ε_{n+1} обновляет и информацию в H_n^ξ .

Предложение 2. Для того, чтобы невырожденная последовательность $\xi = (\xi_n)$ была регулярной, необходимо и достаточно, чтобы нашлись такая обновляющая последовательность $\varepsilon = (\varepsilon_n)$ и числа $(a_k)_{k \geq 0}$ с $\sum_{k=0}^{\infty} |a_k|^2 < \infty$, что (P -п.н.) $\xi_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$.

Предложение 3. Если $\xi = (\xi_n)$ – невырожденная стационарная последовательность, то она допускает разложение Вольда: $\xi_n = \xi_n^s + \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$, где $\sum_{k=0}^{\infty} a_k^2 < \infty$ и $\varepsilon = (\varepsilon_n)$ – некоторая обновляющая последовательность для ξ^r .

Для последовательности $h = (h_n)$ имеем: $\hat{h}_n = \hat{E}(h_n / H_0^h) = \hat{E}(h_n / H_0^\varepsilon) = \hat{h}_n^\varepsilon$. Тогда $\hat{h}_n = \sum_{k=n}^{\infty} a_k \varepsilon_{n-k} = \sum_{k=n}^{\infty} \hat{a}_k^{(n)} h_{n-k}$, где последнее равенство следует из того, что $H_n^\varepsilon = H_n^h$ для всех n .

Приведённая выше формула решает задачу оптимальной линейной экстраполяции значений h_n по прошлой информации $\{h_k, k \leq 0\}$. Но возникают следующие два вопроса:

1. Когда последовательность h_n допускает представление $h_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$.
2. Как находить коэффициенты a_k .

Для того, чтобы выполнялось (1), последовательность h_n должна быть регулярной.

Критерий Колмогорова.

Если $h = h_n$ – стационарная в широком смысле последовательность со спектральной плотностью $f = f(\lambda)$ такой, что $\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty$, то эта последовательность является регулярной.

Приведём графики случайных последовательностей моделей скользящего среднего MA , авторегрессионной модели AR и интегральной модели $ARIMA$.

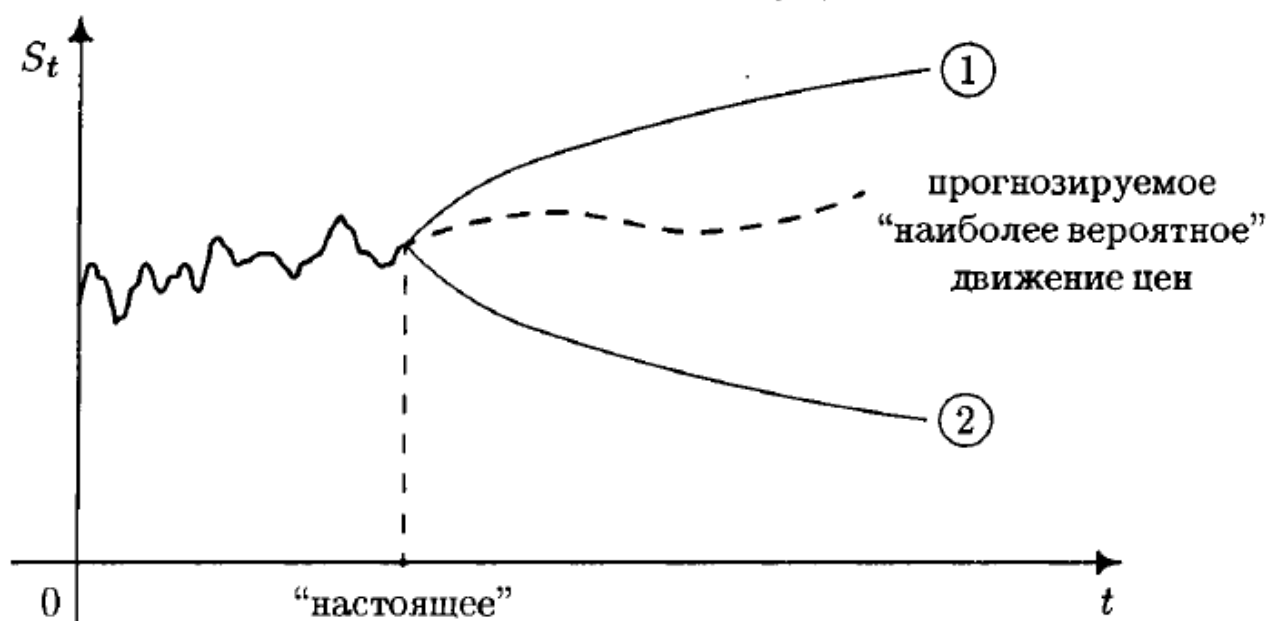


Рис. Область между кривыми 1 и 2 является доверительной областью (с той или иной степенью надежности), в которой будет происходить предполагаемое движение цен в "будущем"

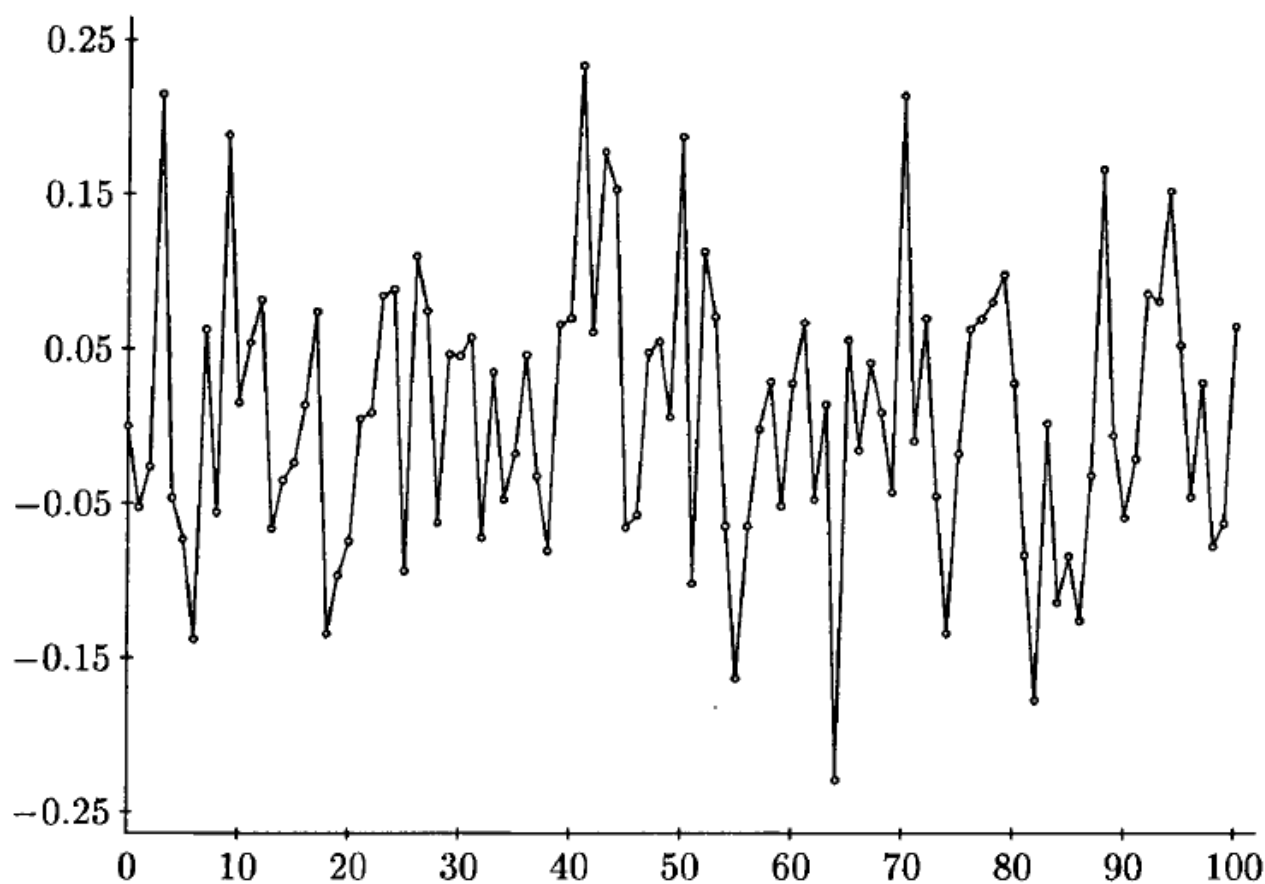


Рис. График компьютерной реализации белого шума $h_n = \sigma \epsilon_n$ с $\sigma = 0.1$ и $\epsilon_n \sim \mathcal{N}(0, 1)$

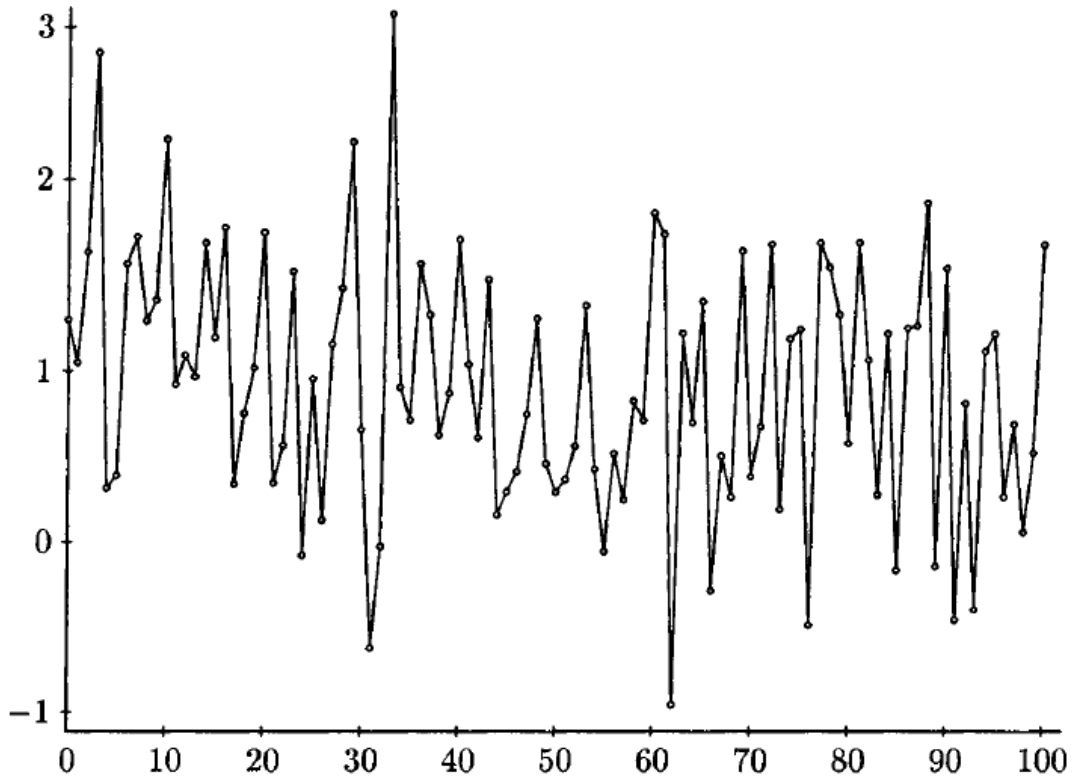


Рис. График компьютерной реализации последовательности $h = (h_n)$, подчиняющейся $MA(1)$ -модели с $h_n = \mu + b_1 \varepsilon_{n-1} + b_0 \varepsilon_n$ с параметрами $\mu = 1, b_1 = 1, b_0 = 0.1$ и $\varepsilon_n \sim \mathcal{N}(0, 1)$

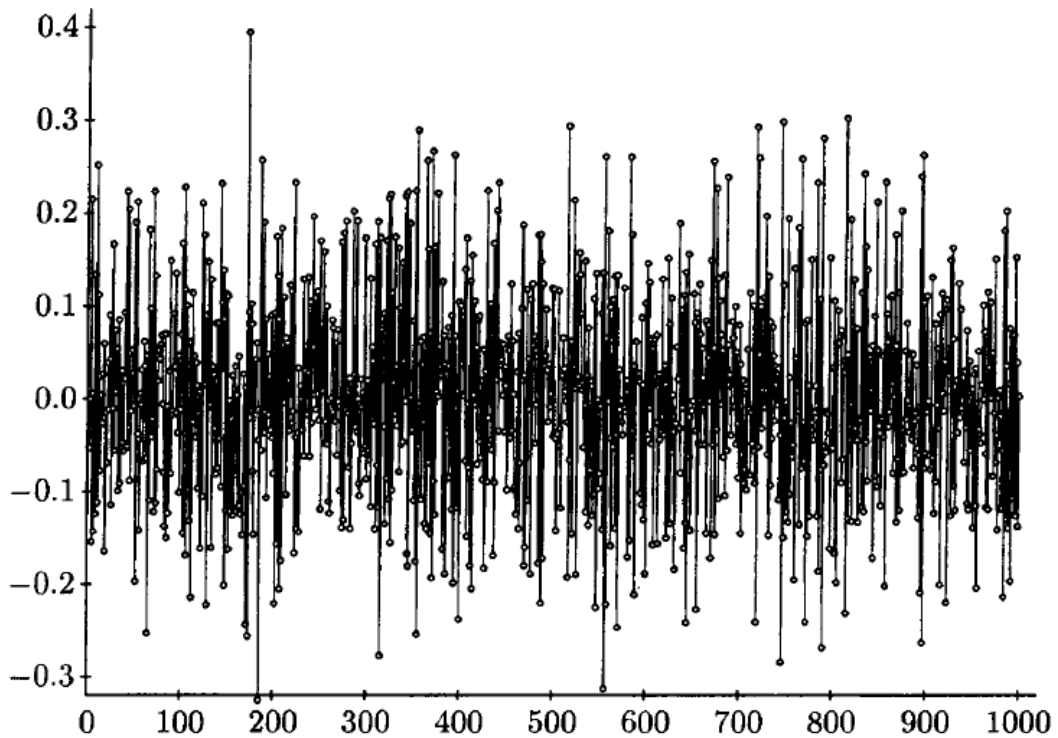


Рис. График компьютерной реализации последовательности $h = (h_n)$, подчиняющейся $AR(2)$ -модели с $h_n = a_0 + a_1 h_{n-1} + a_2 h_{n-2} + \sigma \varepsilon_n$ с параметрами $a_0 = 0, a_1 = -0.5, a_2 = 0.01, \sigma = 0.1; h_0 = h_1 = 0; \varepsilon_n \sim \mathcal{N}(0, 1)$ и $2 \leq n \leq 1000$

Лекция 6

«Модель авторегрессии и скользящего среднего ARMA и интегральная модель ARIMA»

Напомним, что всякая стационарная в широком смысле последовательность $\xi = (\xi_n)$ с $E\xi_n = 0$ допускает спектральное представление

$$\xi_n = \int_{-\pi}^{\pi} e^{i\lambda n} z(d\lambda), \text{ где } z = z(\Delta) \text{ с } \Delta \in B([-\pi, \pi]) \text{ является комплекснозначной}$$

стохастической мерой с ортогональными значениями: $Ez(\Delta_1)\bar{z}(\Delta_2) = 0, \Delta_1 \cap \Delta_2 = \emptyset$.

При этом ковариационная функция $R(n) = \text{cov}(\xi_{k+n}, \xi_k)$ допускает представление $R(n) = \int_{-\pi}^{\pi} e^{i\lambda n} F(d\lambda)$ со спектральной мерой $F(\Delta) = E|z(\Delta)|^2$.

Если существует неотрицательная функция $f = f(\lambda)$ такая, что $F(\Delta) = \int_{\Delta} f(\lambda) d\lambda$, то она называется спектральной плотностью. Сама же функция $F(\lambda) = F((-\infty, \lambda])$ называется спектральной функцией.

Таким образом, если относительно исходной последовательности $h = (h_n)$ известно лишь только то, что она является стационарной в широком смысле со спектральной плотностью $f = f(\lambda)$, удовлетворяющей условию

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty, \text{ то тогда эта последовательность является регулярной.}$$

В интересующих нас линейных моделях последовательность $h = (h_n)$ задавалась не спектральным представлением, а в виде ряда $h_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$, где $\varepsilon = (\varepsilon_n)$ – некоторая ортонормированная последовательность с условиями $E\varepsilon_n = 0, E\varepsilon_i \varepsilon_j = \delta_{ij}$, где δ_{ij} – символ Кронекера: $\delta_{ij} = 1, i = j; \delta_{ij} = 0, i \neq j$.

Последовательность $\varepsilon = (\varepsilon_n)$ является стационарной, её корреляционная (ковариационная) функция $R_{\varepsilon}(n) = \begin{cases} 1, n = 0 \\ 0, n \neq 0 \end{cases}$ и спектральная плотность

$$f_{\varepsilon}(\lambda) = \frac{1}{2\pi}, -\pi \leq \lambda < \pi, \text{ поскольку } R_{\varepsilon}(n) = \int_{-\pi}^{\pi} e^{i\lambda n} \frac{1}{2\pi} d\lambda.$$

Если $h_n = \sum_{k=0}^{\infty} a_k \varepsilon_{n-k}$ и $\sum_{k=0}^{\infty} |a_k|^2 < \infty$, то ковариационная функция последовательности $h = (h_n)$

$$R_h(n) = \int_{-\pi}^{\pi} e^{i\lambda n} f_h(\lambda) d\lambda, f_h(\lambda) = \frac{1}{2\pi} |\Phi(e^{-i\lambda})|^2, \Phi(z) = \sum_{k=0}^{\infty} a_k z^k.$$

Пусть последний ряд имеет радиус сходимости $r > 1$ и не имеет нулей в области $|z| \leq 1$. В этом предположении вопрос о значениях коэффициентов \hat{a}_k оптимального линейного прогноза \hat{h}_n решается следующим образом.

Пусть $\hat{\varphi}_n(\lambda) = e^{i\lambda n} \frac{\Phi_n(e^{-i\lambda})}{\Phi(e^{-i\lambda})}$, $\Phi_n(z) = \sum_{k=n}^{\infty} a_k z^k$. Разложим $\hat{\varphi}_n(\lambda)$ в ряд Фурье:

$\hat{\varphi}_n(\lambda) = \hat{a}_n^{(n)} + \hat{a}_{n+1}^{(n)} e^{-i\lambda} + \hat{a}_{n+2}^{(n)} e^{-2i\lambda} + \dots$. Коэффициенты $\hat{a}_n^{(n)}, \hat{a}_{n+1}^{(n)}, \hat{a}_{n+2}^{(n)}, \dots$ будут в

точности теми, которые определяют линейный прогноз \hat{h}_n величины h_n :
 $\hat{h}_n = \hat{E}(h_n / H_0^h) = \sum_{k=n}^{\infty} \hat{a}_k^{(n)} h_{n-k}$. Таким образом, мы получили возможность построения прогноза значений h_n по прошлым значениям $\{\dots, h_{-1}, h_0\}$, если получится найти соответствующее представление.

Рассмотрим ряд примеров для стационарных моделей $MA(q), AR(p), ARMA(p, q)$, ограничиваясь применением формул прогноза лишь для некоторых значений p и q .

Пример 1. Модель $MA(q)$.

Пусть $h_n = \beta(L)\varepsilon_n$, где $\beta(L) = b_0 + b_1L + \dots + b_qL^q$, то есть $h_n = \sum_{k=0}^q b_k \varepsilon_{n-k}$.

Видим, что $a_k = b_k, 0 \leq k \leq q$ и $a_k = 0, k > q$. Следовательно, $\Phi(z) = \sum_{k=0}^q b_k z^k$ и

$$\Phi_n(z) = \begin{cases} \sum_{k=n}^q b_k z^k, & n \leq q \\ 0, & n > q \end{cases}. \text{ Поэтому для } n \leq q \hat{\varphi}_n(\lambda) = e^{i\lambda n} \frac{\sum_{k=n}^q e^{-i\lambda k} b_k}{\sum_{k=0}^q e^{-i\lambda k} b_k} \text{ и } \hat{\varphi}_n(\lambda) = 0 \text{ для}$$

$n > q$. Следовательно, если $n > q$, то в разложении для $\hat{\varphi}_n(\lambda)$ все коэффициенты $\hat{a}_n = 0$, в этом случае оптимальный прогноз $\hat{h}_n = 0$. Это естественно, так как корреляция между h_n и каждой из величин h_0, h_{-1}, \dots равна 0 при $n > q$.

Если $q = 1$, то $h_n = b_0 \varepsilon_n + b_1 \varepsilon_{n-1}$ и $\hat{\varphi}_1(\lambda) = e^{i\lambda} \frac{e^{-i\lambda} b_1}{b_0 + e^{-i\lambda} b_1}$. Если положить

$b_0 = 1, b_1 = \theta, |\theta| < 1$, то $\Phi(z) = 1 + \theta z$ не имеет нулей в области $|z| \leq 1$. Тогда

$$\hat{\varphi}_1(\lambda) = \frac{\theta}{1 + \theta e^{-i\lambda}} = \theta \sum_{k=0}^{\infty} (-1)^k (\theta e^{-i\lambda})^k = \theta \sum_{k=0}^{\infty} e^{-i\lambda k} (-1)^k \theta^k = \hat{a}_1^{(1)} + \hat{a}_2^{(1)} e^{-i\lambda} + \hat{a}_3^{(1)} e^{-2i\lambda} + \dots$$

Имеем:

$$\hat{a}_1^{(1)} = \theta$$

$$\hat{a}_2^{(1)} = (-1) \cdot \theta^2$$

$$\hat{a}_3^{(1)} = \theta^3$$

...

$$\hat{a}_k^{(1)} = (-1)^{k-1} \cdot \theta^k$$

Таким образом, для модели $MA(1)$ прогноз значения h_1 по прошлому (\dots, h_{-1}, h_0) определяется следующей формулой:

$$\hat{h}_1 = \hat{a}_1^{(1)} h_0 + \hat{a}_2^{(1)} h_{-1} + \dots + \hat{a}_k^{(1)} h_{1-k} + \dots = \theta(h_0 - \theta h_{-1} + \theta^2 h_{-2} + \dots + (-1)^{k-1} \theta^{k-1} h_{1-k} + \dots) =$$

$$= \sum_{k=0}^{\infty} (-1)^k \theta^{k+1} h_{-k}.$$

Отсюда видно, что наибольший вклад в значение прогноза величины h_1 доставляется ближайшим прошлым значением h_0 , и вклад предшествующих значений убывает геометрическим образом (вклад от h_{-k} входит с коэффициентом θ^k , где $|\theta| < 1$).

Пример 2. Модель $AR(p)$.

Пусть для $-\infty < n < \infty$ $h_n = a_1 h_{n-1} + \dots + a_p h_{n-p} + \varepsilon_n$. Стационарное решение данного уравнения имеет вид: $h_n = \sum_{k=0}^{\infty} (c_1 \lambda_1^k + \dots + c_p \lambda_p^k) \varepsilon_{n-k}$, т.е. $a_k = c_1 \lambda_1^k + \dots + c_p \lambda_p^k$. С принципиальной точки зрения прогноз \hat{h}_n по значениям $\{h_k, k \leq 0\}$ даётся формулой $\hat{h}_n = \sum_{k=n}^{\infty} \hat{a}_k^{(n)} h_{n-k}$, где коэффициенты $\hat{a}_k^{(n)}$ получаются из разложения функции $\hat{\varphi}_n(\lambda)$ в ряд Фурье.

В литературе приведены разные приёмы по отысканию формул для прогноза \hat{h}_n . Подсчитаем прогноз для стационарной модели $AR(1): h_n = \theta h_{n-1} + \varepsilon_n, |\theta| < 1$. В этом случае $R(n) = \text{cov}(h_k, h_{k+n})$ определяется формулой $R(n) = \frac{\theta^n}{1 - \theta^2}$. Значит, спектральная функция $f = f(\lambda)$ существует и

$$f(\lambda) = \frac{1}{2\pi} \frac{1}{|1 - \theta e^{-i\lambda}|^2} = \frac{1}{2\pi} |\Phi(e^{-i\lambda})|^2 \Rightarrow \Phi(z) = \frac{1}{1 - \theta z} = \sum_{k=0}^{\infty} (\theta z)^k.$$

$$\text{Отметим, что } f(\lambda) = \frac{1}{2\pi} \left(\sum_{n=0}^{\infty} \frac{\theta^n}{1 - \theta^2} e^{-i\lambda n} + \sum_{n=-\infty}^{-1} \frac{\theta^{-n}}{1 - \theta^2} e^{-i\lambda n} \right).$$

$$\text{Поэтому } \hat{\varphi}_n(\lambda) = e^{i\lambda n} \frac{\Phi_n(e^{-i\lambda})}{\Phi(e^{-i\lambda})} = e^{i\lambda n} \frac{\sum_{k=n}^{\infty} (\theta e^{-i\lambda})^k}{\sum_{k=0}^{\infty} (\theta e^{-i\lambda})^k} = \theta^n.$$

Значит, $\hat{h}_n = \theta^n h_0, n \geq 1$.

Следовательно, для прогнозирования величины h_n по $\{h_k, k \leq 0\}$ достаточно знать лишь значение h_0 , что естественно, если иметь в виду марковский характер последовательности $h = (h_n)$.

Пример 3. Модель $ARMA(p, q)$ с параметрами $p = q = 1$.

В этой модели выражение для $h = (h_n)$ с условием $a_0 = 0$ определяется формулой: $h_n = (a_1 + b_1) \sum_{k=1}^{\infty} a_1^{k-1} \varepsilon_{n-k} + \varepsilon_n$ с $|a_1| < 1$ и $|b_1| < 1$ и, значит, в представлении $h_n = \sum_{k=0}^{\infty} \tilde{a}_k \varepsilon_{n-k}$ коэффициенты $\tilde{a}_0 = 1, \tilde{a}_k = (a_1 + b_1) a_1^{k-1}$.

В рассматриваемом случае функция

$$\Phi(z) = \sum_{k=0}^{\infty} \tilde{a}_k z^k = 1 + \frac{(a_1 + b_1)}{a_1} \sum_{k=1}^{\infty} (a_1 z)^k = 1 + \frac{(a_1 + b_1)}{a_1} \cdot \frac{a_1 z}{1 - a_1 z} \text{ и для } n \geq 1$$

$$\Phi_n(z) = \sum_{k=n}^{\infty} \tilde{a}_k z^k = \frac{(a_1 + b_1)}{a_1} \sum_{k=n}^{\infty} (a_1 z)^k = \frac{(a_1 + b_1)}{a_1} \cdot \frac{(a_1 z)^n}{1 - a_1 z}$$

Поэтому для $z = e^{-i\lambda}$ получаем

$$\hat{\varphi}_n(\lambda) = e^{i\lambda n} \frac{\frac{(a_1 + b_1)}{a_1} \cdot \frac{(a_1 z)^n}{1 - a_1 z}}{1 + \frac{(a_1 + b_1)}{a_1} \cdot \frac{a_1 z}{1 - a_1 z}} = \frac{(a_1 + b_1) a_1^{n-1}}{1 + b_1 z} = (a_1 + b_1) a_1^{n-1} \sum_{k=0}^{\infty} (-1)^k (b_1 z)^k,$$

$$\hat{a}_k^{(n)} = a_1^{n-1} (a_1 + b_1) (-1)^k b_1^k \text{ и}$$

$$\hat{h}_n = a_1^{n-1} (a_1 + b_1) \{h_0 + (-1)b_1 h_{-1} + b_1^2 h_{-2} + \dots\} = a_1^{n-1} (a_1 + b_1) \sum_{k=n}^{\infty} (-1)^k b_1^k h_{n-k}.$$

Модели *EGARCH*, *TGARCH*, *HARCH* и др.

В 1976 году Ф. Блэк подметил в поведении финансовых индексов следующий феномен: отрицательную коррелированность величин h_{n-1} и σ_n , проявляющуюся в том, что эмпирическая ковариация $\text{cov}(h_{n-1}, \sigma_n) < 0$. Этот эффект приводит к тому, что волатильность стремится к возрастанию после падения цен, т.е. после убывания величин возврата. Этот феномен нельзя объяснить в рамках моделей *ARCH*, *GARCH*, поскольку в них волатильность σ_n^2 , будучи зависимой от квадратов предшествующих величин h_{n-1}^2 , нечувствительна к знаку величин h_{n-j} , и значения $h_{n-j} = \Delta$ и $h_{n-j} = -\Delta$ приводят в *GARCH*-моделях к одному и тому же значению будущей волатильности σ_n^2 .

Для объяснения обнаруженного Ф. Блэком эффекта Д. Нельсон в 1990 году предложил так называемую модель *Exponential GARCH* (p, q), или *EGARCH* (p, q), в которой учёт асимметрии осуществлялся тем, что вместо величин $h_{n-i}^2 = \sigma_{n-i}^2 \varepsilon_{n-i}^2$, входящих в *GARCH*-модель, вводится линейная комбинация величин ε_{n-i} и $|\varepsilon_{n-i}|$. Именно, снова считается, что $h_n = \sigma_n \varepsilon_n$, но для σ_n предполагается выполненным следующее соотношение:

$$\ln \sigma_n^2 = \alpha_0 + \sum_{i=1}^p \alpha_i [\theta \varepsilon_{n-i} + \gamma (|\varepsilon_{n-i}| - \sqrt{2/\pi})] + \sum_{j=1}^q \beta_j \ln \sigma_{n-j}^2.$$

Заметим, что $E|\varepsilon_{n-i}| = \sqrt{2/\pi}$. Поскольку $h_{n-i} = \sigma_{n-i} \varepsilon_{n-i}$ и $\sigma_{n-i} \geq 0$, то знаки у h_{n-i} и ε_{n-i} совпадают. Следовательно, если $\varepsilon_{n-i} = \Delta > 0$, то соответствующий вклад в σ_n^2 определяется величиной $\Delta(\theta + \gamma)$. Если $\varepsilon_{n-i} = -\Delta < 0$, то вклад будет равен $\Delta(-\theta + \gamma)$. В общем, *EGARCH*-модели сохраняют свойства моделей *GARCH*, но позволяют ухватывать эффект асимметрии.

Примером другой разновидности может служить модель *Threshold GARCH* (p, q), или *TGARCH* (p, q).

Рассмотрим модель *Threshold AR*, или *TAR*:

$h_n = \sum_{i=1}^k I_{A_i}(h_{n-d}) (\alpha_0^i + \alpha_1^i h_{n-1} + \dots + \alpha_p^i h_{n-p})$, где d – параметр запаздывания, A_1, \dots, A_k

– непересекающиеся множества в R такие, что $\sum_{i=1}^k A_i = R$.

$$\text{Например, пусть } h_n = \begin{cases} \alpha_0^1 + \alpha_1^1 h_{n-1} + \alpha_2^1 h_{n-2}, h_{n-2} > 0 \\ \alpha_0^2 + \alpha_1^2 h_{n-1} + \alpha_2^2 h_{n-2}, h_{n-2} \leq 0 \end{cases}.$$

По определению последовательность $h = (h_n)$ описывается $TGARCH(p, q)$ -моделью, если $h_n = \sigma_n \varepsilon_n$, где

$$\sigma_n = a_0 + \sum_{i=1}^p [a_i h_{n-i}^+ + b_i h_{n-i}^-] + \sum_{j=1}^q [c_j \sigma_{n-j}^+ + d_j \sigma_{n-j}^-] \text{ и } x^+ = \max(x, 0), x^- = -\min(x, 0). \text{ В}$$

этой модели не предполагается неотрицательность коэффициентов и неотрицательность волатильностей σ_n , хотя смысл σ_n^2 как условной дисперсии $E(h_n^2 / F_{n-1}^h)$ сохраняет свою силу.

Поскольку

$$h_n = \sigma_n \varepsilon_n = (\sigma_n^+ - \sigma_n^-) (\varepsilon_n^+ - \varepsilon_n^-) = [\sigma_n^+ \varepsilon_n^+ + \sigma_n^- \varepsilon_n^-] - [\sigma_n^- \varepsilon_n^+ + \sigma_n^+ \varepsilon_n^-], \text{ то}$$

$$h_n^+ = [\sigma_n^+ \varepsilon_n^+ + \sigma_n^- \varepsilon_n^-] \text{ и } h_n^- = \sigma_n^- \varepsilon_n^+ + \sigma_n^+ \varepsilon_n^-.$$

Эти соотношения дают возможность переписать выражение для волатильности в следующем виде: $\sigma_n = \alpha_0 + \sum_{i=1}^{p^*} \alpha_i(\varepsilon_{n-i}) \sigma_{n-i}^+ + \sum_{i=1}^{q^*} \beta_i(\varepsilon_{n-i}) \sigma_{n-i}^-$ со значениями $p^* = \max(p, q)$ и функциями $\alpha_i(\varepsilon_{n-i})$ и $\beta_i(\varepsilon_{n-i})$, являющимися линейными комбинациями величин ε_{n-i}^- и ε_{n-i}^+ .

Есть определённая техническая сложность изучения таких моделей, заключающаяся в отсутствии марковского свойства. Однако в простых случаях, например, для $p = q = 1$, можно провести достаточно полное исследование свойств этих моделей.

$$\text{Действительно, при } p = q = 1 \text{ } \sigma_n = a_0 + [a_1 h_{n-1}^+ + b_1 h_{n-1}^-] + [c_1 \sigma_{n-1}^+ + d_1 \sigma_{n-1}^-], \text{ или}$$

$$\sigma_n = \alpha_0 + \alpha_1(\varepsilon_{n-1}) \sigma_{n-1}^+ + \beta_1(\varepsilon_{n-1}) \sigma_{n-1}^-, \text{ где } \alpha_0 = a_0, \alpha_1(\varepsilon_{n-1}) = a_1 \varepsilon_{n-1}^+ + b_1 \varepsilon_{n-1}^- + c_1,$$

$$\beta_1(\varepsilon_{n-1}) = a_1 \varepsilon_{n-1}^- + b_1 \varepsilon_{n-1}^+ + d_1.$$

Если $\alpha_0 = 0$, то $\sigma_n^+ = (\alpha_1(\varepsilon_{n-1}))^+ \sigma_{n-1}^+ + (\beta_1(\varepsilon_{n-1}))^+ \sigma_{n-1}^-$, $\sigma_n^- = (\alpha_1(\varepsilon_{n-1}))^- \sigma_{n-1}^+ + (\beta_1(\varepsilon_{n-1}))^- \sigma_{n-1}^-$. То есть, последовательность $(\sigma_n^+, \sigma_n^-, \varepsilon_n)_{n \geq 1}$ относительно потока (F_n) является марковской, что даёт возможность проводить исследование её свойств обычными марковскими методами.

Остановимся ещё на одном феномене – эффекте долгой памяти, или сильного последействия в эволюции цен $S = (S_n)_{n \geq 0}$.

Существуют разные меры зависимости случайных последовательностей от прошлого: коэффициенты эргодичности, перемешивания и др. Например, можно измерять степень пропадания зависимости от прошлого у действительной стационарной последовательности $X = (X_n)$ тем, как быстро

величина $\sup_A E|P(X_{n+m} \in A | X_1, \dots, X_n) - P(X_{n+m} \in A)| \rightarrow 0$ при $m \rightarrow \infty$, где \sup берётся по всем борелевским множествам $A \subseteq R$.

Стандартной же мерой зависимости является автокорреляционная функция.

Отметим, что, согласно многолетним статистическим исследованиям, у временных финансовых рядов обнаруживается более сильная корреляционная зависимость в поведении последовательностей $|h| = (h_n)_{n \geq 1}$ и $h^2 = (h_n^2)_{n \geq 1}$, чем та, которая получается в случае моделей *ARCH, GARCH* и, тем более, *MA, AR, ARMA*.

Напомним, что для *ARCH(1)* $\rho(k) = \text{Corr}(h_{n-k}^2, h_n^2) = \alpha_1^k, \alpha_1 < 1$. Видно, что моделях *ARCH(1), GARCH(1,1)* корреляция убывает геометрическим образом к нулю, то есть память быстро забывает прошлое.

Часто принято говорить, что стационарная в широком смысле последовательность $Y = (Y_n)$ имеет долгую память, или сильное последствие, если её автокорреляционная функция $\rho(k)$ имеет гиперболический характер убывания к нулю: $\rho(k) \sim ck^{-\rho}, k \rightarrow \infty$ при некотором $\rho > 0$.

Таким характером убывания автокорреляционной функции обладает, например, фрактальный гауссовский шум $Y = (Y_n)_{n \geq 1}$ со значениями $Y_n = X_n - X_{n-1}$, где $X = (X_t)_{t \geq 0}$ является фрактальным броуновским движением с параметром $0 < H < 1$. Для этого движения

$$\text{cov}(X_s, X_t) = \frac{1}{2} \{ |t|^{2H} + |s|^{2H} - |t-s|^{2H} \} EX_1^2, \text{cov}(Y_n, Y_{n+k}) = \frac{\sigma^2}{2} \{ |k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H} \},$$

где $\sigma^2 = DY_n$ и, следовательно, автокорреляционная функция $\rho(k) = \text{cov}(Y_n, Y_{n+k})$ имеет при $k \rightarrow \infty$ гиперболический характер: $\rho(k) \sim H(2H-1)k^{2H-2}$. Заметим,

что при $\frac{1}{2} < H < 1$ $\sum_{k=1}^{\infty} \rho(k) = \infty$. При $H = \frac{1}{2}$ (обычное броуновское движение) величины $Y = (Y_n)$ образуют гауссовский белый шум с $\rho(k) = 0, k \geq 1$. Если же $0 < H < \frac{1}{2}$, то $\sum_{k=1}^{\infty} |\rho(k)| < \infty, \sum_{k=1}^{\infty} \rho(k) = 0$.

В модели *HARCH(p)* характер убывания автокорреляционной функции для модулей и квадратов величин h_n более медленный, чем для моделей *ARCH, GARCH*. Этот же эффект долгой памяти свойственен и моделям *FIGARCH*.

Модель *HARCH(p)* порядка p задаётся следующим образом: $h_n = \sigma_n \varepsilon_n$, где $\sigma_n^2 = \alpha_0 + \sum_{j=1}^p \alpha_j \left(\sum_{i=1}^j h_{n-i} \right)^2$ с $\alpha_0 > 0, \alpha_p > 0, \alpha_j \geq 0, j = 1, \dots, p-1$. В частности, при $p=1$ *HARCH(1) = ARCH(1)*, то есть $\sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2$. В случае $p=2$ $\sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2 + \alpha_2 (h_{n-1} + h_{n-2})^2$.

Приведём некоторые свойства этой модели.

Прежде всего, отметим, что наличие члена $(h_{n-1} + h_{n-2})^2$ даёт возможность также ухватывать эффекты асимметрии.

Далее, если $\alpha_0 + \alpha_1 + \alpha_2 < 1$, то существует стационарное значение

$$Eh_n^2 = E\sigma_n^2 = \frac{\alpha_0}{1 - \alpha_1 - 2\alpha_2}.$$

Аналогичным образом, рассматривая $E\sigma_n^4$ и используя то, что $Eh_{n-1}h_{n-2} = Eh_{n-1}^3h_{n-2} = Eh_{n-1}h_{n-2}^3 = 0$, находим, что при $(\alpha_1 + \alpha_2)^2 + \alpha_2^2 < \frac{1}{3}$ существует

$$\text{стационарное значение } Eh_n^4 = \frac{C}{\frac{1}{3} - (\alpha_1 + \alpha_2)^2 - \alpha_2^2},$$

$$C = \frac{\alpha_0^2 [1 + 2\alpha_2(\alpha_1 + 3\alpha_2) - (\alpha_1 + 2\alpha_2)^2]}{[1 - (\alpha_1 + 2\alpha_2)]^2}. \text{ Отметим, что } E\sigma_n^4 = \frac{1}{3}Eh_n^4.$$

Найдём теперь значения автокорреляционной функции (для h_n^2).

Пусть $R(k) = Eh_n^2 h_{n-k}^2$. Тогда для $k = 1$ в предположении стационарности

$$\begin{aligned} R(1) &= E\sigma_n^2 \varepsilon_n^2 h_{n-1}^2 = E\sigma_n^2 h_{n-1}^2 = E(h_{n-1}^2 [\alpha_0 + \alpha_1 h_{n-1}^2 + \alpha_2 (h_{n-1} + h_{n-2})^2]) = \\ &= \alpha_0 Eh_{n-1}^2 + \alpha_1 Eh_{n-1}^4 + \alpha_2 Eh_{n-1}^4 + 2\alpha_2 Eh_{n-1}^3 h_{n-2} + \alpha_2 Eh_{n-1}^2 h_{n-2}^2. \end{aligned}$$

$$\text{Если } \alpha_2 < 1, \text{ то } R(1) = \frac{\alpha_0 Eh_{n-1}^2 + (\alpha_1 + \alpha_2) Eh_{n-1}^4}{1 - \alpha_2}.$$

Далее,

$$\begin{aligned} R(k) &= Eh_n^2 h_{n-k}^2 = E\sigma_n^2 h_{n-k}^2 = E[\alpha_0 + \alpha_1 h_{n-1}^2 + \alpha_2 (h_{n-1} + h_{n-2})^2] h_{n-k}^2 = \\ &= \alpha_0 Eh_{n-k}^2 + (\alpha_1 + \alpha_2) R(k-1) + \alpha_2 R(k-2), \end{aligned}$$

где $R(0) = Eh_n^4$.

Отсюда ясно, что в стационарном случае автокорреляционная функция $\rho(k) = \text{corr}(h_n^2, h_{n-k}^2)$ подчиняется уравнению ($k \geq 2$):

$$\rho(k) = A + B\rho(k-1) + C\rho(k-2), A = \frac{\alpha_0 Eh_n^2}{Dh_n^2}, B = \frac{(\alpha_1 + \alpha_2)(Eh_n^2)^2}{Dh_n^2}, C = \frac{(\alpha_2 - 1)(Eh_n^2)^2}{Dh_n^2}$$

$$\text{и } \rho(0) = 1, \rho(1) = \frac{R(1) - (Eh_n^2)^2}{Dh_n^2}.$$

Описание моделей семейства *GARCH* давалось в рамках условного подхода. При этом условное распределение $\text{Law}(h_n / \sigma_n)$ было всегда нормальным $N(0, \sigma_n^2)$, где σ_n^2 предсказуемым образом зависели от прошлого. Один из таких естественно возникающих при таком подходе вопросов заключается в следующем: каковы безусловные распределения $\text{Law}(h_n)$,

$\text{Law}(h_1, \dots, h_n), n \geq 1$?

Рассмотрим следующую модель.

Пусть $h_n = \sigma_n \varepsilon_n$, где (ε_n) – стандартная гауссовская последовательность, а

$\sigma_n^2 = a\sigma_{n-1}^2 + b\delta_n$, (δ_n) – последовательность неотрицательных независимых устойчивых случайных величин с индексом устойчивости $0 < \alpha < 1$. Последовательности (ε_n) и (δ_n) независимы.

Если предположить, что $0 \leq a < 1$, то последовательными итерациями находим, что $\sigma_n^2 = b \sum_{k=0}^{\infty} a^k \delta_{n-k} + \lim_{m \rightarrow \infty} a^m \sigma_{n-m-1}^2$. В силу свойств автомодельности устойчивых распределений $\sum_{k=0}^{\infty} a^k \delta_{n-k} \stackrel{d}{=} \left(\sum_{k=0}^{\infty} a^{ak} \right)^{1/\alpha} \delta_1$, и в предположениях $0 < \alpha < 1, 0 \leq a < 1$ уравнение имеет (конечное) неотрицательное стационарное решение (σ_n^2) : $\sigma_n^2 \stackrel{d}{=} b \left(\frac{1}{1-a^\alpha} \right)^{1/\alpha} \delta_1$.

Отсюда и из определения $h_n = \sigma_n \varepsilon_n$ заключаем, что стационарное одномерное распределение $\text{Law}(h_n)$ является устойчивым с индексом устойчивости, равным 2α .

В заключение остановимся на эффекте тяжелых хвостов, встречающимся в нелинейных стохастических моделях.

Рассмотрим $ARCH(1)$ -модель $h = (h_n)_{n \geq 0}$ с $h_n = \sqrt{\alpha_0 + \alpha_1 h_{n-1}^2} \varepsilon_n$ для $n \geq 1$, начальным условием h_0 , не зависящим от стандартной гауссовской последовательности $\varepsilon = (\varepsilon_n)_{n \geq 1}$ с $\alpha_0 > 0, 0 < \alpha_1 < 1$.

Оказывается, что при подходящем выборе распределения величины h_0 рассматриваемой модели существует решение $h = (h_n)_{n \geq 0}$, являющееся строго стационарным процессом, для которого при достаточно малом $\alpha_1 > 0$ имеет место эффект тяжёлых хвостов: $P(h_n > x) \sim cx^{-\gamma}$, где $c > 0, \gamma > 0$.

Также отметим, что в том случае, когда в исходной модели неизвестны параметры $\theta = (a_0, a_1, c)$, для отыскания оценок $\hat{\sigma}_n$ параметра σ_n по h_1, \dots, h_n часто используется байесовский метод, основанный на предположении наличия априорного распределения $\pi(\theta)$ для θ . Тогда возможно нахождение апостериорного распределения $\pi(\theta, \sigma_n / h_1, \dots, h_n)$, из которого затем можно найти апостериорные распределения $\pi(\theta / h_1, \dots, h_n)$ и $\pi(\sigma_n / h_1, \dots, h_n)$, дающие возможность построения оценок $\hat{\theta}_n$ и $\hat{\sigma}_n$, например, как апостериорных средних или тех значений, где апостериорные плотности обращаются в максимум.

Удивительным является то обстоятельство, что чисто детерминистические динамические системы могут проявлять свойства стохастического белого шума. Поэтому интересно было бы рассмотреть модели динамического хаоса. Это нелинейные хаотические модели, а также проблематика различимости хаотических и стохастических последовательностей (для выяснения истинной природы нерегулярности финансовых данных). Один из возможных подходов может быть основан на понятии сложности для индивидуальных реализаций.

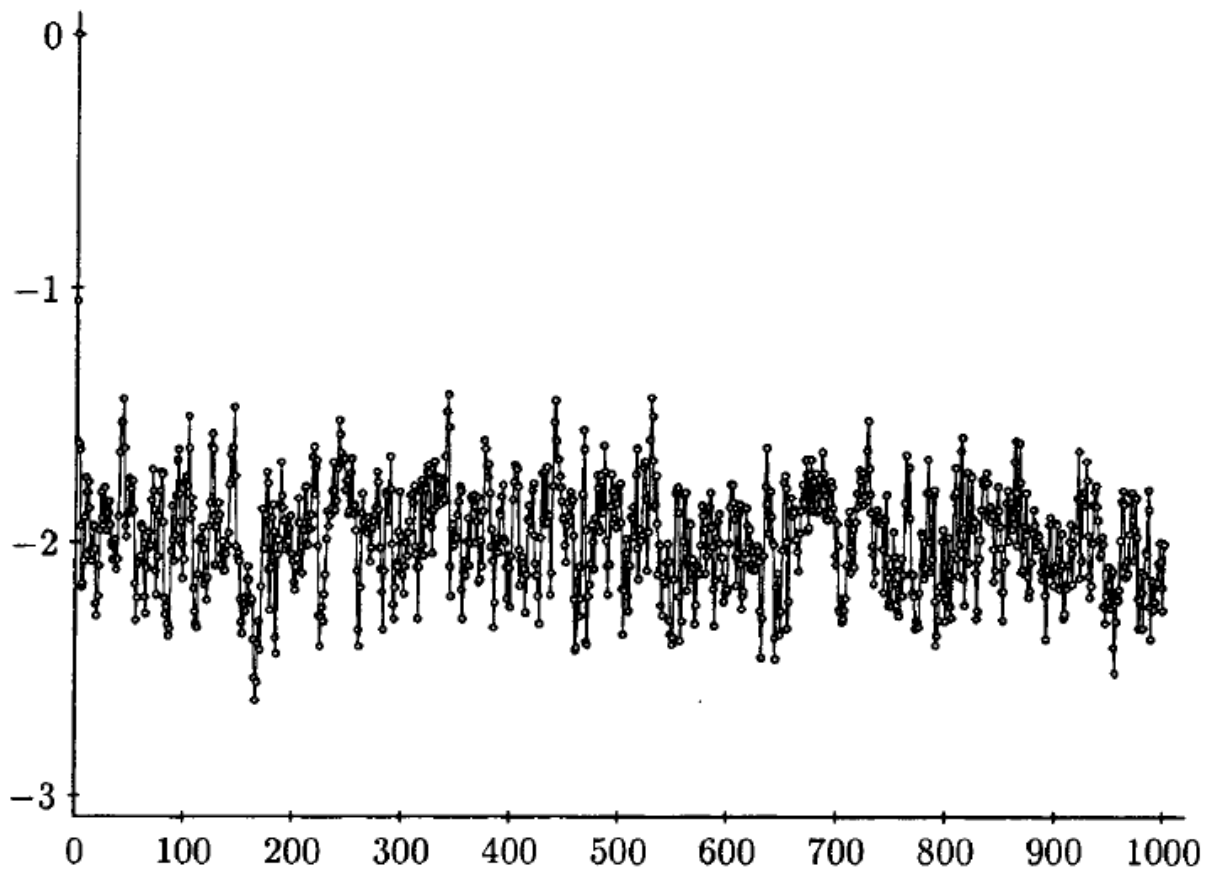


Рис. График компьютерной реализации последовательности $h = (h_n)$, подчиняющейся $ARMA(1, 1)$ -модели с $h_n = a_0 + a_1 h_{n-1} + b_1 \varepsilon_{n-1} + \sigma \varepsilon_n$ с параметрами $a_0 = -1$, $a_1 = 0.5$, $b_1 = 0.1$, $\sigma = 0.1$; $h_0 = 0$ и $0 \leq n \leq 1000$

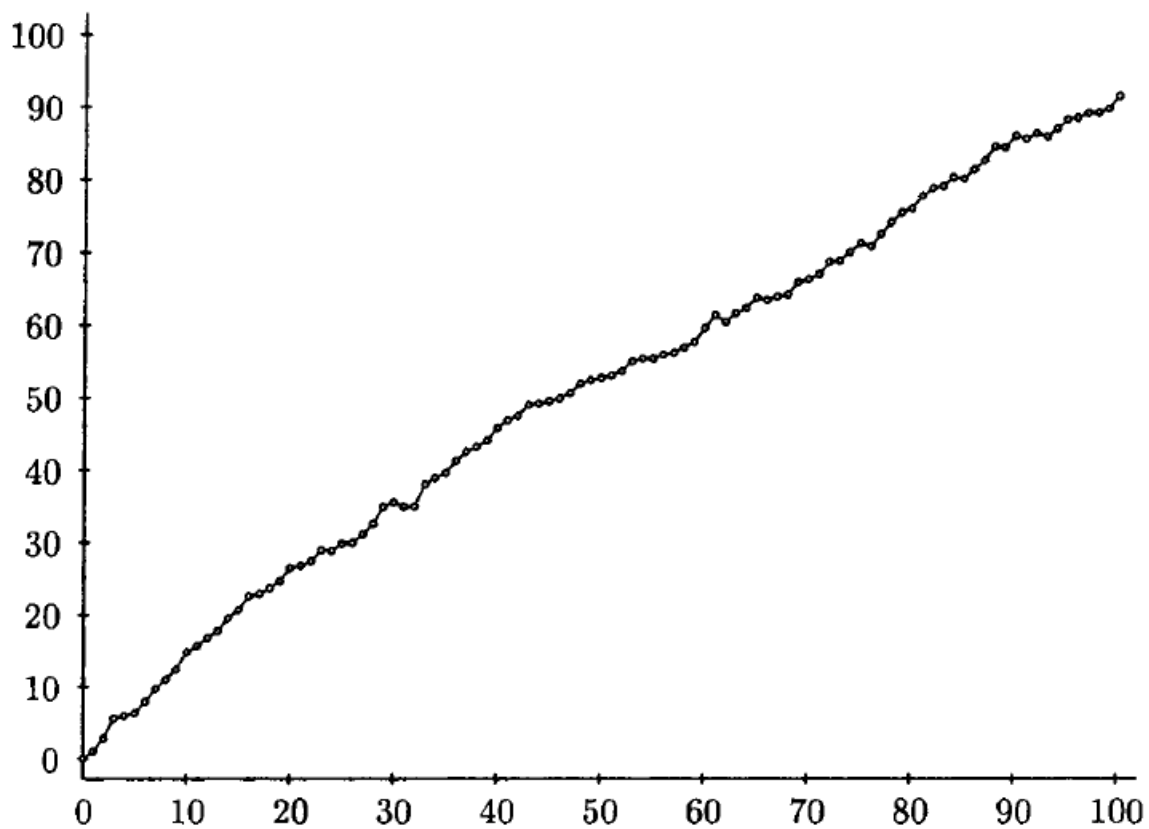


Рис. График компьютерной реализации последовательности $x = (x_n)$, подчиняющейся $ARIMA(0, 1, 1)$ -модели с $\Delta x_n = \mu + b_1 \varepsilon_{n-1} + b_0 \varepsilon_n$ с параметрами $\mu = 1, b_1 = 1, b_0 = 0.1; x_0 = 0$

Лекция 9

Нелинейные стохастические условно-гауссовские модели *ARCH* и *GARCH*

Рассмотрим вопрос о предсказании будущего движения цен, считая, что последовательность $h = (h_n)$ подчиняется модели *ARCH*(p).

Поскольку последовательность $h = (h_n)$ является мартингал-разностью, то $E(h_{n+m} / F_n) = 0$, и оптимальная в среднеквадратическом смысле оценка $\hat{h}_{n+m} = E(h_{n+m} / F_n^h) = E(E(h_{n+m} / F_n) / F_n^h) = 0$. Тривиальность этой оценки делает целесообразным рассмотрение вопроса о предсказании будущих значений от нелинейных функций от h_{n+m} , например, величин $h_{n+m}^2, |h_{n+m}|$.

$$\begin{aligned} & \text{Имеем } (F_n^h = \sigma(h_1, \dots, h_n)): \hat{h}_{n+m}^2 = \\ & = E(h_{n+m}^2 / F_n^h) = E(\sigma_{n+m}^2 \varepsilon_{n+m}^2 / F_n^h) = E(E(\sigma_{n+m}^2 \varepsilon_{n+m}^2 / F_{n+m-1}^\varepsilon) / F_n^h) = E(\sigma_{n+m}^2 / F_n^h) = \hat{\sigma}_{n+m}^2. \end{aligned}$$

Отсюда ясно, что вопрос предсказания будущих значений h_{n+m}^2 сводится к вопросу предсказания волатильности σ_{n+m}^2 по результатам прошлых наблюдений h_0, h_1, \dots, h_n .

$$\begin{aligned} & \text{Поскольку } \sigma_{n+m}^2 = \alpha_0 + \alpha_1 \sigma_{n+m-1}^2 \varepsilon_{n+m-1}^2, \text{ то по индукции находим, что} \\ & \sigma_{n+m}^2 = \alpha_0 + \alpha_1 [\alpha_0 + \alpha_1 \sigma_{n+m-2}^2 \varepsilon_{n+m-2}^2] \varepsilon_{n+m-1}^2 = \dots = \alpha_0 + \alpha_0 \prod_{j=1}^{m-1} \prod_{i=1}^j \alpha_1 \varepsilon_{n+j-i+1}^2 + \sigma_n^2 \prod_{i=1}^m \alpha_1 \varepsilon_{n+m-i}^2. \end{aligned}$$

Отсюда, беря условное математическое ожидание $E(\cdot / F_n^h)$ и учитывая независимость в величинах последовательности (ε_n) , находим, что

$$\hat{h}_{n+m}^2 = \hat{\sigma}_{n+m}^2 = E(\sigma_{n+m}^2 / F_n^h) = \alpha_0 + \alpha_0 \sum_{j=1}^{m-1} \alpha_1^j + h_n^2 \alpha_1^m = \alpha_0 \frac{1 - \alpha_1^m}{1 - \alpha_1} + \alpha_1^m h_n^2.$$

Как и следовало ожидать, при $m \rightarrow \infty$ оценки \hat{h}_{n+m}^2 сходятся (с вероятностью единица) к стационарному значению $Eh_n^2 = \frac{\alpha_0}{1 - \alpha_1}$.

Напомним, что для нормального распределения $N(\mu, \sigma)$ интервалы $(\mu - \sigma, \mu + \sigma)$ и $(\mu - 1.65\sigma, \mu + 1.65\sigma)$ являются (с определённой степенью точности) доверительными интервалами с уровнями надёжности 68% и 90% соответственно.

Поскольку $S_{n+m} = S_n e^{h_{n+1} + \dots + h_{n+m}}$ и $E[(h_{n+1} + \dots + h_{n+m})^2 / F_n^h] = E(h_{n+1}^2 / F_n^h) + \dots + E(h_{n+m}^2 / F_n^h) = \hat{\sigma}_{n+1}^2 + \dots + \hat{\sigma}_{n+m}^2$, то в первом приближении в качестве доверительных интервалов (с уровнями значимости 68% и 90%) можно взять интервалы $(S_n e^{-\sqrt{\hat{\sigma}_{n+1}^2 + \dots + \hat{\sigma}_{n+m}^2}}, S_n e^{\sqrt{\hat{\sigma}_{n+1}^2 + \dots + \hat{\sigma}_{n+m}^2}})$ и $(S_n e^{-\sqrt{\hat{\sigma}_{n+1}^2 + \dots + \hat{\sigma}_{n+m}^2}}, S_n e^{\sqrt{\hat{\sigma}_{n+1}^2 + \dots + \hat{\sigma}_{n+m}^2}})$, соответственно.

При этом слова «в первом приближении» связаны с тем, что на самом деле величины h_k не являются нормально распределёнными.

Успех условно-гауссовской модели *ARCH*(p), давшей объяснение таких феноменов в поведении финансовых индексов, как кластерность, тяжёлые хвосты, вытянутость плотности распределения величин h_n и многих других,

породил целый ряд различных её обобщений, преследующих цель дать возможные объяснения ряда других эффектов, обнаруживаемых методами статистического анализа.

Одним из первых обобщений модели $ARCH(p)$ явилась введённая в 1986 г. Т.Боллерселем Обобщённая (Generalized) $ARCH$ -модель, характеризуемая двумя параметрами p и q и обозначаемая $GARCH(p, q)$.

В этой модели, как и для $ARCH(p)$ -модели, снова берётся $h_n = \sigma_n \varepsilon_n$, но относительно формирования волатильности σ_n предполагается, что $\sigma_n^2 = \alpha_0 + \sum_{i=1}^p \alpha_i h_{n-i}^2 + \sum_{j=1}^q \beta_j \sigma_{n-j}^2$, где $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$. Если все $\beta_j = 0$, то получаем модель $ARCH(p)$.

Основным преимуществом $GARCH(p, q)$ -моделей сравнительно с $ARCH(p)$ -моделью является то, что при подгонке статистических данных моделями $ARCH(p)$ часто приходится обращаться к слишком большим значениям p , в то время как при подгонке $GARCH(p, q)$ -моделями можно ограничиваться лишь небольшими p и q . Но это экспериментальный факт.

Анализ моделей $GARCH(p, q)$, в которых волатильность σ_n предполагается зависящей (предсказуемым образом) как от $h_{n-i}^2, i \leq p$, так и от $\sigma_{n-j}^2, j \leq q$, проводится аналогично анализу $ARCH(p)$ -моделей.

Приведём ряд простых формул, относящихся к модели $GARCH(1, 1)$.

$h_n = \sigma_n \varepsilon_n, \sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2 + \beta_1 \sigma_{n-1}^2$, где $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$. Отсюда ясно, что $Eh_n^2 = \alpha_0 + (\alpha_1 + \beta_1)Eh_{n-1}^2$, и стационарное значение Eh_n^2 существует при $\alpha_1 + \beta_1 < 1$ и задаётся формулой $Eh_n^2 \equiv \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$. Если $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$, то

существует стационарное значение $Eh_n^4 \equiv \frac{3\alpha_0^2(1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)}$. Тем

самым, стационарный коэффициент эксцесса

$K \equiv \frac{Eh_n^4}{(Eh_n^2)^2} - 3 = \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2}$. Нетрудно найти и стационарное значение

автокорреляционной функции $\rho(k)$: $\rho(1) = \frac{\alpha_1(1 - \alpha_1\beta_1 - \beta_1^2)}{1 - 2\alpha_1\beta_1 - \beta_1^2}$,

$\rho(k) = (\alpha_1 + \beta_1)^{k-1} \rho(1), k > 1$.

Также отметим, что $\hat{h}_{n+m}^2 = \hat{\sigma}_{n+m}^2 = E(\sigma_{n+m}^2 / F_n^h) = \alpha_0 \frac{1 - \gamma^m}{1 - \gamma} + \gamma^{m-1}(\alpha_1 h_n^2 + \beta_1 \sigma_n^2)$

где $\gamma = \alpha_1 + \beta_1$.

Модели семейства $ARCH$, эволюционирующие в дискретном времени, имеют соответствующие аналоги и в непрерывном времени. Более того, при подходящей нормировке можно получить (слабую) сходимость решений

стохастических разностных уравнений, определяющих ARCH, GARCH-модели, к решениям соответствующих стохастических дифференциальных уравнений.

Для определённости рассмотрим следующую модификацию GARCH(1,1)-модели (GARCH(1,1)-M-модель).

Пусть Δ – временной шаг, $H^{(\Delta)} = (H_{k\Delta}^{(\Delta)}), k = 0, 1, \dots$, где $H_{k\Delta}^{(\Delta)} = H_0^{(\Delta)} + h_{\Delta}^{(\Delta)} + \dots + h_{k\Delta}^{(\Delta)}$, $h_{k\Delta}^{(\Delta)} = c(\sigma_{k\Delta}^{(\Delta)})^2 + (\sigma_{k\Delta}^{(\Delta)})\varepsilon_{k\Delta}$, причём $\varepsilon_{k\Delta} \sim N(0, \Delta)$, c – константа и $(\sigma_{k\Delta}^{(\Delta)})^2 = \alpha_0(\Delta) + (\sigma_{(k-1)\Delta}^{(\Delta)})^2(\beta(\Delta) + \alpha_1(\Delta)\varepsilon_{(k-1)\Delta}^2)$. Будем считать, что заданы начальные условия $H_0^{(\Delta)} = H_0, \sigma_0^{(\Delta)} = \sigma_0$ при всех $\Delta > 0$, где (H_0, σ_0) – пара случайных величин, не зависящая от гауссовских последовательностей случайных величин.

Вложим последовательность $(H^{(\Delta)}, \sigma^{(\Delta)}) = (H_{k\Delta}^{(\Delta)}, \sigma_{k\Delta}^{(\Delta)})_{k \geq 0}$ в схему с непрерывным временем $t \geq 0$, полагая $H_t^{(\Delta)} = H_{k\Delta}^{(\Delta)}, \sigma_t^{(\Delta)} = \sigma_{k\Delta}^{(\Delta)}$ при $k\Delta \leq t < (k+1)\Delta$.

Из общих результатов теории слабой сходимости случайных процессов следует, что при некоторых условиях на входящие коэффициенты последовательность процессов $(H^{(\Delta)}, \sigma^{(\Delta)})$ слабо сходится (в пространстве Скорохода) к некоторому диффузионному процессу $(H, \sigma) = (H_t, \sigma_t)_{t \geq 0}$.

При $\alpha_0(\Delta) = \alpha_0\Delta, \alpha_1(\Delta) = \alpha\sqrt{\frac{\Delta}{2}}, \beta(\Delta) = 1 - \alpha\sqrt{\frac{\Delta}{2}} - \beta\Delta$ предельным при $\Delta \rightarrow 0$ процессом (H, σ) является процесс, подчиняющийся следующим стохастическим дифференциальным уравнениям:

$dH_t = c\sigma_t^2 dt + \sigma_t dW_t^{(1)}, d\sigma_t^2 = (\alpha_0 - \beta\sigma_t^2)dt + \alpha\sigma_t^2 dW_t^{(2)}$, где $(W^{(1)}, W^{(2)})$ – два независимых стандартных броуновских движения, независимых также от начальных значений $(H_0, \sigma_0) \equiv (H_0^{(\Delta)}, \sigma_0^{(\Delta)})$.

До сих пор при описании эволюции последовательностей $h = (h_n)$, где $h_n = \ln \frac{S_n}{S_{n-1}}$ и S_n – значение цены в момент времени n , мы исходили из той гипотезы, что эти величины имеют стохастическую природу, т.е. $S_n = S_n(\omega), h_n = h_n(\omega)$ являются случайными величинами, заданными на некотором фильтрованном вероятностном пространстве $(\Omega, F, (F_n)_{n \geq 1}, P)$, моделирующими стохастическую неопределённость состояний природы.

С другой стороны, достаточно хорошо известно, что даже совсем простые нелинейные детерминистические системы типа $x_{n+1} = f(x_n; \lambda)$ или $x_{n+1} = f(x_n, x_{n-1}, \dots, x_{n-k}; \lambda)$, где λ – некоторый параметр, могут порождать (при соответствующих начальных условиях) последовательности (x_0, x_1, x_2, \dots) , поведение которых весьма схоже с поведением стохастических последовательностей.

Возникает вопрос о том, не являются ли многие экономические и финансовые ряды в действительности не стохастическими, а хаотическими.

Пусть $x = (x_n)$ – хаотическая последовательность, порождённая некоторой динамической системой с распределением вероятностей $F = F(x)$ для x_0 , являющимся инвариантным для данной системы.

Пусть теперь $\tilde{x} = (\tilde{x}_n)$ – стохастическая последовательность, состоящая из независимых одинаково распределённых величин с одномерным распределением $F = F(x)$.

Образуем величины $M_n = \max(x_0, x_1, \dots, x_n)$ и $\tilde{M}_n = \max(\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_n)$ и пусть $F_n(x) = P(M_n \leq x)$, $\tilde{F}_n(x) = P(\tilde{M}_n \leq x)$.

Идея подхода основана на том наблюдении, что максимум служит хорошей характеристикой, позволяющей улавливать отличия стохастических последовательностей от хаотических.

В теории предельных теорем для экстремальных значений хорошо известны необходимые и достаточные условия, при которых величины $a_n(\tilde{M}_n - b_n)$, $n \geq 1$ имеют нетривиальное предельное распределение $\lim_n P(a_n(\tilde{M}_n - b_n) \leq x) = \tilde{G}(x)$ для некоторых констант $a_n > 0$ и b_n , $n \geq 1$.

Приведём некоторые примеры.

Если $F(x) = 1 - x^{-\rho}$, $x \geq 1$, $\rho > 0$, то $P\left(\frac{\tilde{M}_n}{n^{1/\rho}} \leq x\right) \rightarrow \exp(-x^{-\rho})$, $x > 0$.

Если $F(x) = 1 - (-x)^\rho$, $-1 \leq x \leq 0$, $\rho > 0$, то для $x < 0$ $P(n^{1/\rho} \tilde{M}_n \leq x) \rightarrow \exp(-|x|^\rho)$.

Если $F(x) = 1 - e^{-x}$, $x \geq 0$, то для $x \in R$ $P(\tilde{M}_n - \ln n \leq x) \rightarrow \exp(-e^{-x})$.

Если $F(x) = \Phi(x)$ – стандартное нормальное распределение, то $P((2 \ln n)^{1/2}(\tilde{M}_n - b_n) \leq x) \rightarrow (-e^{-x})$, $x \in R$, где b_n выбраны так, что $P(x_0 > b_n) = \frac{1}{2}$.

В этом случае $b_n \sim (2 \ln n)^{1/2}$.

Для преобразования Бернулли инвариантное распределение $F(x) = x$, $x \in (0,1)$, и при $a_n = n$, $b_n = 1 - n^{-1}$ предельное распределение есть $\tilde{G}(x) = \exp(x-1)$, $x \leq 1$.

Для $F(x) = 1 - p^2(x)$, где $p(x) = \frac{1-x}{2}$, $x \in (-1,1)$ и при $a_n = \sqrt{n}$, $b_n = 1 - 2/\sqrt{n}$ находим, что $\tilde{G}(x) = \exp\left(-\left(\frac{x}{2} - 1\right)^2\right)$.

Для $F(x) = \frac{2}{\pi} \arcsin \sqrt{x}$ при соответствующей нормализации можно найти, что $\tilde{G}(x) = \exp(-(1-2x)^{1/2})$.

Можно привести ещё много примеров, посвящённых проблематике различимости хаотических и стохастических последовательностей. При этом критерии различимости могут быть различными.

Подведём итог изучения авторегрессионной модели условной неоднородности $ARCH(p)$. Итак, мы предполагаем, что последовательность $\varepsilon = (\varepsilon_n)_{n \geq 1}$ является единственным источником случайности, $F_n = \sigma(\varepsilon_1, \dots, \varepsilon_n)$, $\mu_n = E(h_n / F_{n-1}) = 0$ и $\sigma_n^2 = E(h_n^2 / F_{n-1}) = \alpha_0 + \sum_{i=1}^p \alpha_i h_{n-i}^2$, где $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, p, F_0 = \{\emptyset, \Omega\}$ и h_{1-p}, \dots, h_0 – заданные начальные константы.

Иначе говоря, условная дисперсия σ_n^2 является функцией от значений $h_{n-1}^2, \dots, h_{n-p}^2$.

Эта модель, как уже было отмечено, была введена в 1982 году Р.Энглеом и названа $ARCH(p)$ (Autoregressive Conditional Heteroskedastic Model – Авторегрессионная модель условной неоднородности). Она оказалась весьма удачной при объяснении ряда нетривиальных свойств временных финансовых рядов таких, как, например, эффект кластерности (скупенности) значений величин h_n .

Итак, $h_n = \sigma_n \varepsilon_n$, где $\varepsilon = (\varepsilon_n)$ – последовательность независимых нормально распределённых случайных величин, $\varepsilon_n \sim N(0,1)$, а σ_n^2 определены выше.

Если $\mu_n = a_0 + a_1 h_{n-1} + \dots + a_r h_{n-r}$, то $h_n = a_0 + a_1 h_{n-1} + \dots + a_r h_{n-r} + \sigma_n \varepsilon_n$. Эти модели иногда обозначают $AR(r)/ARCH(p)$.

Пусть $Eh_n^2 < \infty$. Положим $v_n = h_n^2 - \sigma_n^2$. Тогда $h_n^2 = \alpha_0 + \sum_{i=1}^p \alpha_i h_{n-i}^2 + v_n$, где $E(v_n / F_{n-1}) = E(h_n^2 / F_{n-1}) - \sigma_n^2 = 0$, то есть последовательность $v = (v_n)$ образует мартингал-разность.

Таким образом, $ARCH(p)$ -модель может рассматриваться как авторегрессионная модель $AR(p)$ для последовательности (h_n^2) с шумом $v = (v_n)$, являющимся мартингал-разностью.

Рассмотрим обобщённую авторегрессионную модель условной неоднородности $GARCH(p, q)$ (Generalized Autoregressive Conditional Heteroskedastic Model), введённую в 1986 году Т.Боллерслевом.

Пусть $\mu_n = E(h_n / F_{n-1}) = 0$ и $\sigma_n^2 = E(h_n^2 / F_{n-1}) = \alpha_0 + \sum_{i=1}^p \alpha_i h_{n-i}^2 + \sum_{j=1}^q \beta_j \sigma_{n-j}^2$ с $\alpha_0 > 0, \alpha_i, \beta_j \geq 0$ и начальными условиями $(h_{1-p}, \dots, h_0), (\sigma_{1-q}^2, \dots, \sigma_0^2)$, которые для простоты можно считать константами.

Модель $GARCH(p, q)$ – это последовательность $h = (h_n)$, $h_n = \sigma_n \varepsilon_n$, где $(\varepsilon_1, \varepsilon_2, \dots)$ – последовательность независимых одинаково распределённых стандартных нормальных случайных величин.

Будем обозначать $\alpha(L)h_{n-1}^2 = \sum_{i=1}^p \alpha_i h_{n-i}^2$, где L – оператор сдвига ($L^i h_{n-1}^2 = h_{(n-1)-i}^2$) и $\beta(L)\sigma_{n-1}^2 = \sum_{j=1}^q \beta_j \sigma_{n-j}^2$. В этих обозначениях

$\sigma_n^2 = \alpha_0 + \alpha(L)h_{n-1}^2 + \beta(L)\sigma_{n-1}^2$. Если, как и выше, положить $v_n = h_n^2 - \sigma_n^2$, то получим

$$\begin{aligned} h_n^2 &= v_n + \sigma_n^2 = v_n + \alpha_0 + \alpha(L)h_{n-1}^2 + \beta(L)(h_{n-1}^2 - v_{n-1}) = \\ &= \alpha_0 + (\alpha(L) + \beta(L))h_{n-1}^2 - \beta(L)v_{n-1} + v_n. \end{aligned}$$

Иначе говоря, $h_n^2 = \alpha_0 + (\alpha(L) + \beta(L))h_{n-1}^2 + v_n - \beta(L)v_{n-1}$. Тем самым, $GARCH(p, q)$ -модель можно рассматривать как модель авторегрессии скользящего среднего $ARMA(\max(p, q), q)$, для последовательности (h_n^2) с шумом (v_n) , который является мартингал-разностью.

В частности, для модели $ARCH(1)$ с $h_n = \sigma_n \varepsilon_n$, $\sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2$ находим, полагая $v_n = h_n^2 - \sigma_n^2$, что $h_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2 + v_n$, где шум (v_n) образует мартингал-разность.

Разнообразные обобщения $ARCH$ и $GARCH$ моделей, такие, как $ARCH$, $ARCH$, $ARCH$, $ARCH$, $ARCH$, $ARCH$, ... связаны в конечном счёте с той или иной спецификацией величин $\sigma_n^2 = E(h_n^2 / F_{n-1})$ как функций, измеримых относительно σ -алгебр $F_{n-1} = \sigma(\varepsilon_1, \dots, \varepsilon_{n-1})$.

Рассмотрим модель стохастической волатильности.

Во всех предыдущих моделях источник случайности был один. Он задавался гауссовской последовательностью независимых величин $\varepsilon = (\varepsilon_n)$. Модели стохастической волатильности включают в себя два источника случайности: $\varepsilon = (\varepsilon_n)$ и $\delta = (\delta_n)$, которые в простейшем случае предполагаются независимыми и стандартными гауссовскими последовательностями, то есть состоящими из независимых, $N(0,1)$ -распределённых случайных величин.

Пусть $G_n = \sigma(\delta_1, \dots, \delta_n)$. Положим $h_n = \sigma_n \varepsilon_n$, где σ_n являются G_n -измеримыми.

Тогда ясно, что $Law(h_n / G_n) = N(0, \sigma_n^2)$, то есть G_n -условное распределение h_n является гауссовским с параметрами 0 и σ_n^2 .

Положим $\sigma_n = \exp(\Delta_n / 2)$. Тогда $\sigma_n^2 = \exp(\Delta_n)$, где Δ_n — G_n -измеримы.

Весьма популярны модели, где последовательность (Δ_n) является авторегрессионной моделью, $(\Delta_n) \in AR(p)$, $\Delta_n = a_0 + a_1 \Delta_{n-1} + \dots + a_p \Delta_{n-p} + c \delta_n$.

Естественным обобщением является схема, в которой $h_n = \mu_n + \sigma_n \varepsilon_n$, где μ_n и σ_n — G_n -измеримы.

В том случае, когда последовательность $\varepsilon = (\varepsilon_n)$ образует нормально распределённую стационарную последовательность с $E\varepsilon_n = 0$, $E\varepsilon_n^2 = 1$, а $\sigma = (\sigma_n)$ не зависят от $\varepsilon = (\varepsilon_n)$, соответствующая модель носит название модели Тейлора.

В общем, целый ряд гауссовских и условно-гауссовских моделей находят применение в финансовой математике и финансовой инженерии. При этом рассматриваются различные их обобщения, которые лучше отражают реальные значения финансовых индексов.

Концепция эффективного рынка обосновывает гипотезу мартингалности (нормированных) цен, делая тем самым понятие мартингала одним из основных при исследовании динамики эволюции цен как стохастических последовательностей или процессов с определёнными свойствами их распределений. Однако, при проведении конкретных расчётов одного лишь знания мартингалности распределений слишком мало, нужна более подробная структура этих распределений, что приводит к необходимости детального рассмотрения разных вероятностно-статистических моделей с целью выявления тех из них, свойства распределений которых лучше всего согласуются со свойствами эмпирических распределений, построенных по статистическим данным. Именно поэтому имеет смысл рассматривать модели, позволяющие объяснять свойства, обнаруживаемые при анализе статистических данных, например, образованных временными финансовыми рядами.

С точки зрения теоретического анализа, наиболее привлекательным является предположение гауссовости распределений $\text{Law}(h_1, \dots, h_n)$ величин h_1, \dots, h_n . Но статистическая обработка данных многих финансовых рядов показывает, что предположение гауссовости не всегда адекватно отражает истинную картину поведения цен.

Если пытаться искать альтернативу предположению о гауссовости безусловных распределений $\text{Law}(h_1, \dots, h_n)$ последовательности $h = (h_n)_{n \geq 1}$, то, имея в виду разложение Дуба, которое определяется с привлечением условных математических ожиданий $E(h_n / F_{n-1})$, вполне естественно предположить, что не безусловные, а условные распределения вероятностей $\text{Law}(h_n / F_{n-1})$ являются гауссовскими: $\text{Law}(h_n / F_{n-1}) = N(\mu_n, \sigma_n^2)$ с некоторыми F_{n-1} -измеримыми $\mu_n = \mu_n(\omega)$ и $\sigma_n^2 = \sigma_n^2(\omega)$. Во избежание рассмотрения тривиальных случаев, требующих специальных оговорок, будем считать, что при всех n и ω величины $\sigma_n(\omega) \neq 0$.

Итак, регулярное условное распределение $P(h_n \leq x / F_{n-1})$ задаётся формулой (для всех $x \in R$ и $\omega \in \Omega$):

$$P(h_n \leq x / F_{n-1})(\omega) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{-\infty}^x \exp\left(-\frac{(y - \mu_n(\omega))^2}{2\sigma_n^2(\omega)}\right) dy.$$

Из регулярности этого условного распределения вытекает, что $E(h_n / F_{n-1})(\omega) = \int_{-\infty}^{+\infty} x dP(h_n \leq x / F_{n-1})(\omega)$, что в рассматриваемом случае приводит к формуле $E(h_n / F_{n-1}) = \mu_n$.

Аналогично $D(h_n / F_{n-1}) = \sigma_n^2$.

Таким образом, параметры μ_n и σ_n^2 имеют традиционный смысл – это условное среднее и условная дисперсия условного распределения $\text{Law}(h_n / F_{n-1})$.

Само же распределение $\text{Law}(h_n)$ является взвесью (смесью) условных гауссовских распределений $\text{Law}(h_n / F_{n-1})$ с усреднением по распределению величин μ_n и σ_n^2 .

Отметим, что класс распределений, образованных взвесью нормальных распределений $N(\mu, \sigma^2)$ со случайными параметрами $\mu = \mu(\omega), \sigma^2 = \sigma^2(\omega)$ является весьма широким. Существует много разных частных случаев таких распределений.

Наряду с последовательностью $h = (h_n)$, целесообразно ввести стандартную условно-гауссовскую последовательность $\varepsilon = (\varepsilon_n)_{n \geq 1}$ F_n -измеримых случайных величин таких, что $\text{Law}(\varepsilon_n / F_{n-1}) = N(0, 1)$, где $F_0 = \{\emptyset, \Omega\}$.

Понятно, что эта последовательность является мартингал-разностью, поскольку $E(\varepsilon_n / F_{n-1}) = 0$. Более того, это последовательность независимых случайных величин, имеющих стандартное нормальное распределение $N(0, 1)$, поскольку $\text{Law}(\varepsilon_n / \varepsilon_1, \dots, \varepsilon_{n-1}) = N(0, 1)$.

В силу сделанного выше предположения $\sigma_n(\omega) \neq 0 (n \geq 1, \omega \in \Omega)$ величины $\varepsilon_n, n \geq 1$, определяемые формулой $\varepsilon_n \equiv (h_n - \mu_n) / \sigma_n$, образуют стандартную гауссовскую последовательность. Следовательно, можно считать, что рассматриваемые условно-гауссовские (относительно потока (F_n) и вероятности P) последовательности $h = (h_n)_{n \geq 1}$ представимы в виде: $h_n = \mu_n + \sigma_n \varepsilon_n$, где $\varepsilon = (\varepsilon_n)$ – последовательность независимых F_n -измеримых случайных величин, имеющих стандартное нормальное распределение, $N(0, 1)$. По поводу представления последовательности $h = (h_n)_{n \geq 1}$ в общем случае, то есть когда σ_n может обращаться в ноль, существует много книг и статей.

Ясно, что более детальное изучение вероятностных свойств последовательностей $h = (h_n)$ и $S = (S_n)$ зависит от конкретизации структуры величин μ_n и σ_n^2 .

Заметим, что с точки зрения распределений последовательности $h = (h_n)$ и желая иметь условную гауссовость целесообразно бывает рассмотрение этого свойства в следующем контексте.

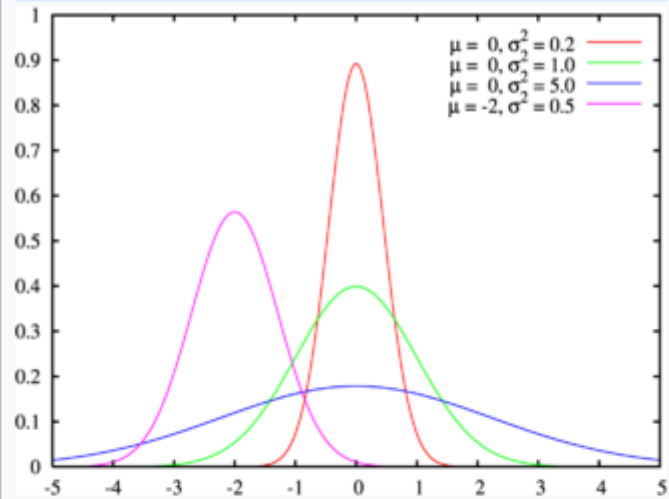
Пусть (G_n) – подфльтрация (F_n) , то есть $G_n \subseteq F_n, G_n \subseteq G_{n+1}$. Например, $G_n = F_{n-1}$.

Предположим, что $\text{Law}(h_n / G_n) = N(\mu_n, \sigma_n^2)$ с $\mu_n = E(h_n / G_n)$, $\sigma_n^2 = D(h_n / G_n)$. В этом случае распределение $\text{Law}(h_n)$ также является смесью гауссовских.

В линейных и нелинейных стохастических гауссовских и условно-гауссовских моделях для $n \geq 1$ специфицируются значения μ_n, σ_n и задаются начальные условия (\dots, h_{-1}, h_0) и $(\dots, \varepsilon_{-1}, \varepsilon_0)$ для h, ε .

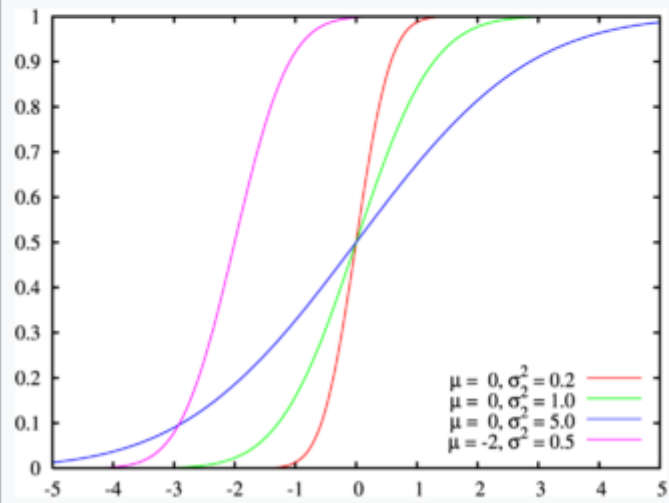
Приведём некоторые графики в качестве приложения к лекции.

Нормальное распределение



Зеленая линия соответствует стандартному нормальному распределению

Плотность вероятности



Цвета на этом графике соответствуют графику сверху

Функция распределения

Обозначение $N(\mu, \sigma^2)$

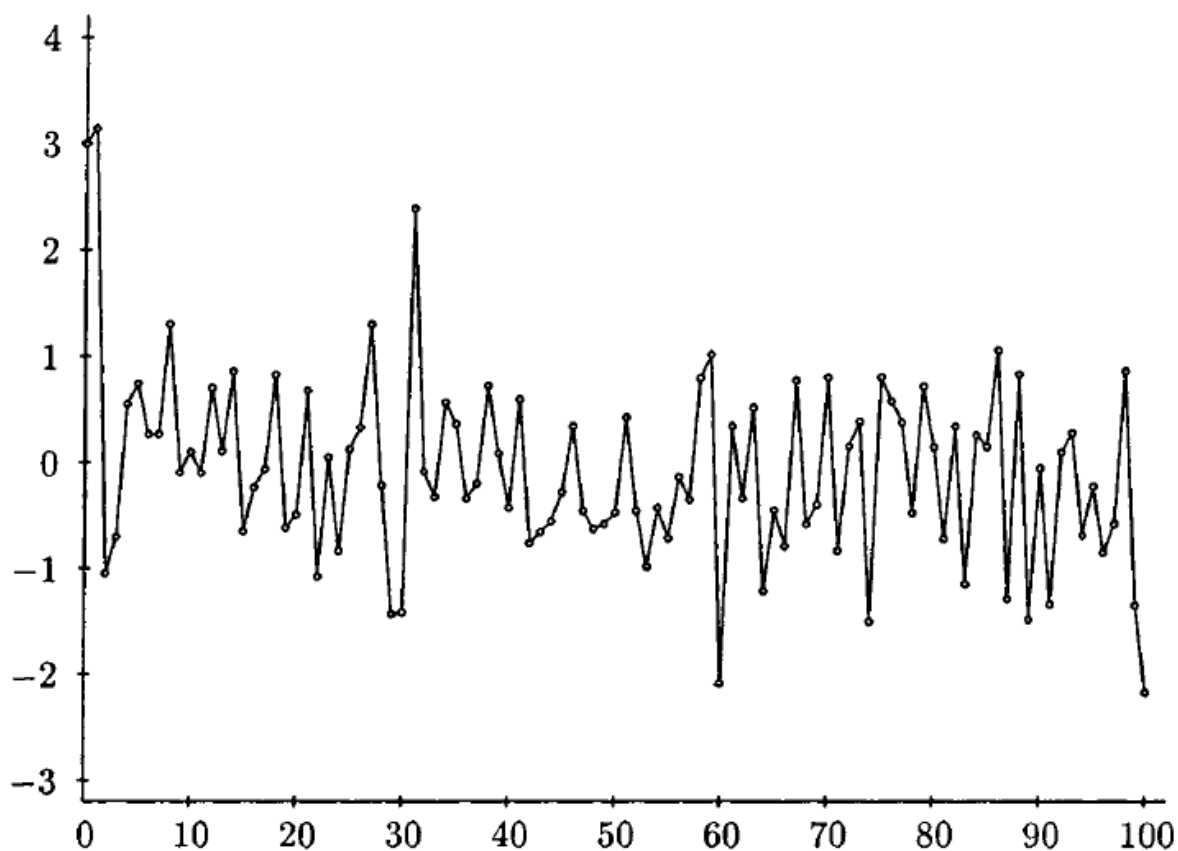


Рис. График компьютерной реализации последовательности $h = (h_n)$, подчиняющейся $ARCH(1)$ -модели с $h_n = \sqrt{\alpha_0 + \alpha_1 h_{n-1}^2} \varepsilon_n$ с параметрами $\alpha_0 = 0.9, \alpha_1 = 0.2; h_0 = 3$ и $0 \leq n \leq 100$

Лекция 10

Модели стохастической волатильности

Эти модели характеризуются наличием двух источников случайности $\varepsilon = (\varepsilon_n)$ и $\delta = (\delta_n)$, определяющих поведение последовательности $h = (h_n)$ со значениями $h_n = \sigma_n \varepsilon_n$, где $\sigma_n = e^{\frac{\Delta_n}{2}}$, последовательность $(\Delta_n) \in AR(p)$:

$\Delta_n = a_0 + \sum_{i=1}^p a_i \Delta_{n-i} + c \delta_n$. Обе последовательности $\varepsilon = (\varepsilon_n)$ и $\delta = (\delta_n)$ будем предполагать независимыми стандартными гауссовскими, при этом будем говорить, что $h = (h_n)$ подчиняется $SV(p)$ -модели, то есть модели стохастической волатильности.

Рассмотрим свойства этой модели, предполагая $p=1$ и $|a_1| < 1$: $h_n = \sigma_n \varepsilon_n$, $\ln \sigma_n^2 = a_0 + a_1 \ln \sigma_{n-1}^2 + c \delta_n$. Пусть $F_n^{\varepsilon, \delta} = \sigma(\varepsilon_1, \dots, \varepsilon_n; \delta_1, \dots, \delta_n)$, $F_n^\delta = \sigma(\delta_1, \dots, \delta_n)$, $F_n^\varepsilon = \sigma(\varepsilon_1, \dots, \varepsilon_n)$. Тогда $E(h_n / F_n^\delta) = \sigma_n E \varepsilon_n = 0$ и $E(h_n / F_{n-1}^{\varepsilon, \delta}) = E(\sigma_n \varepsilon_n / F_{n-1}^{\varepsilon, \delta}) = E(\sigma_n E(\varepsilon_n / F_{n-1}^{\varepsilon, \delta}) \vee \sigma(\delta_n)) / F_{n-1}^{\varepsilon, \delta}) = E(\sigma_n E(\varepsilon_n / F_{n-1}^{\varepsilon, \delta})) = 0$, поскольку $E \varepsilon_n = 0$.

Тем самым, относительно потока $(F_n^{\varepsilon, \delta})$ величины $h = (h_n)$ образуют мартингал-разность. Но не относительно (F_n^δ) , поскольку h_n не являются F_n^δ -измеримыми.

Далее, $E h_n^2 = E \sigma_n^2 E \varepsilon_n^2 = E \sigma_n^2 = E e^{\Delta_n}$. Будем предполагать, что $\Delta_0 \sim N\left(\frac{a_0}{1-a_1}, \frac{c^2}{1-a_1^2}\right)$. Тогда последовательность $\Delta = (\Delta_n)$ удовлетворяет авторегрессионной схеме $AR(1)$: $\Delta_n = a_0 + a_1 \Delta_{n-1} + c \delta_n$ и является стационарной последовательностью.

Имеем: $E h_n^2 = E e^{\Delta_n} = e^{\frac{a_0}{1-a_1}} e^{\frac{c^2}{2(1-a_1^2)}}$, где для подсчёта $E e^{\Delta_n}$ мы воспользовались тем, что для случайной величины $\xi \sim N(0,1)$ и для любого σ $E e^{\sigma \xi - \frac{\sigma^2}{2}} = 1$. Аналогичным образом находим, что $E|h_n| = E|\varepsilon_n| E \sigma_n = \sqrt{\frac{2}{\pi}} E^{\frac{\Delta_n}{2}} = \sqrt{\frac{2}{\pi}} e^{\frac{a_0}{2(1-a_1)}} e^{\frac{c^2}{8(1-a_1^2)}}$.

Рассмотрим теперь ковариационные свойства последовательностей $h = (h_n)$ и $h^2 = (h_n^2)$.

Имеем: $E h_n h_{n+1} = 0$. Вообще, для любого $k \geq 1$ $E h_n h_{n+k} = 0$. Тем самым, последовательность $h = (h_n)$ состоит из некоррелированных случайных величин. Если $R_h(k) = E h_n h_{n+k}$, то $R_h(k) = \begin{cases} E h_n^2, & k = 0 \\ 0, & k > 0 \end{cases}$. Далее,

$$\begin{aligned} E h_n^2 h_{n-1}^2 &= E \sigma_n^2 \sigma_{n-1}^2 = E e^{\Delta_n + \Delta_{n-1}} = E(e^{\Delta_{n-1}} E(e^{a_0 + a_1 \Delta_{n-1} + c \delta_n} | \Delta_{n-1})) = \\ &= E e^{a_0 + \Delta_{n-1}(1+a_1)} E e^{c \delta_n} = e^{a_0 + \frac{c^2}{2}} E e^{(1+a_1)\Delta_{n-1}} = e^{a_0 + \frac{c^2}{2}} e^{\frac{a_0(1+a_1)}{1-a_1}} E e^{(1+a_1)\left(\Delta_{n-1} - \frac{a_0}{1-a_1}\right)} = \end{aligned}$$

$$= e^{\frac{2a_0+c^2}{1-a_1} \frac{1+a_1}{2} \frac{c^2}{1-a_1^2}} = e^{\frac{2a_0+c^2}{1-a_1} \frac{c^2}{2} \frac{1+a_1}{1-a_1}} = e^{\frac{2a_0+c^2}{1-a_1} \frac{2}{1-a_1}} = e^{\frac{2a_0+c^2}{1-a_1}}.$$

Поэтому

$$\text{cov}(h_n^2, h_{n-1}^2) = e^{\frac{2a_0+c^2}{1-a_1}} - e^{\frac{2a_0}{1-a_1}} e^{\frac{c^2}{1-a_1^2}} = e^{\frac{2a_0}{1-a_1}} e^{\frac{c^2}{1-a_1^2}} \left(e^{\frac{c^2 a_1}{1-a_1^2}} - 1 \right).$$

Как и следовало ожидать, величины h_n^2 и h_{n-1}^2 положительно коррелированы в случае $a_1 > 0$ и отрицательно коррелированы при $a_1 < 0$.

Наряду с полученными формулами $E\sigma_n^2 = e^{\frac{a_0}{1-a_1}} e^{\frac{c^2}{2(1-a_1^2)}}$, $E\sigma_n^2 \sigma_{n-1}^2 = e^{\frac{2a_0}{1-a_1}} e^{\frac{c^2}{1-a_1}}$, приведём и более общие формулы $E\sigma_n^r = e^{\frac{ra_0}{2(1-a_1)} + \frac{r^2}{8} \frac{c^2}{1-a_1^2}}$, $E\sigma_n^r \sigma_{n-1}^s = E\sigma_n^r \sigma_n^s e^{\frac{rsc^2}{4(1-a_1^2)}}$ для положительных r и s .

Эти формулы можно использовать для подсчёта различных моментов от

$$h = (h_n). \text{ Например, } E|h_n| = \sqrt{\frac{2}{\pi}} E\sigma_n = \sqrt{\frac{2}{\pi}} e^{\frac{a_0}{2(1-a_1)} + \frac{c^2}{8(1-a_1^2)}}, Eh_n^4 = 3E\sigma_n^4,$$

$$Eh_n^2 h_{n-k}^2 = E\sigma_n^2 \sigma_{n-k}^2, E|h_n h_{n-k}| = \frac{2}{\pi} E\sigma_n \sigma_{n-k}.$$

Следовательно, стационарный коэффициент эксцесса:

$$K = \frac{Eh_n^4}{(Eh_n^2)^2} - 3 = 3 \left(\frac{E\sigma_n^4 - (E\sigma_n^2)^2}{(E\sigma_n^2)^2} \right) = 3 \frac{D\sigma_n^2}{(E\sigma_n^2)^2} \geq 0.$$

Это говорит о том, что модели стохастической волатильности с двумя источниками случайности $\varepsilon = (\varepsilon_n)$ и $\delta = (\delta_n)$, так же, как и модели семейства ARCH, позволяют описывать последовательности $h = (h_n)$, у которых плотности распределения вероятностей величин h_n имеют вытянутость в окрестности среднего значения $Eh_n = 0$.

Остановимся на вопросах построения оценок $\hat{\sigma}_n$ волатильности по результатам наблюдений h_1, \dots, h_n . Если $h_n = \mu + \sigma_n \varepsilon_n$, то $Eh_n = \mu$ и $E|h_n - \mu| =$

$$= E|\sigma_n \varepsilon_n| = \sqrt{\frac{2}{\pi}} E\sigma_n \Rightarrow \hat{\sigma}_n = \sqrt{\frac{\pi}{2}} |h_n - \bar{h}_n|, \text{ где } \bar{h}_n = \frac{1}{n} \sum_{k=1}^n h_k, \text{ если } \mu \text{ неизвестно, и}$$

$$\hat{\sigma}_n = \sqrt{\frac{\pi}{2}} |h_n - \mu|, \text{ если } \mu \text{ известно.}$$

Другой метод оценивания σ_n^2 исходит из того факта, что $Eh_n^2 = E\sigma_n^2$, то есть основан на свойствах моментов второго порядка.

В качестве оценки для σ_n^2 можно взять оценку $\hat{\sigma}_n^2 = h_n^2$. Она будет несмещённой, однако её среднеквадратическая ошибка $E|\hat{\sigma}_n^2 - \sigma_n^2|^2 =$

$$= E|h_n^2 - \sigma_n^2|^2 = Eh_n^4 - 2Eh_n^2 \sigma_n^2 + E\sigma_n^4 = 4E\sigma_n^4 - 2E\sigma_n^4 = 2E\sigma_n^4 = 2 \exp \left\{ \frac{2a_0}{1-a_1} + \frac{2c^2}{1-a_1^2} \right\}$$

может оказаться довольно большой.

Естественно, что если величины $\sigma_k^2, k \leq n$ являются коррелированными, то можно пытаться при конструировании оценок σ_n^2 использовать не только одно наблюдение h_n^2 , но и предшествующие наблюдения $h_{n-1}^2, h_{n-2}^2, \dots$. При этом, конечно, понятно, что если величины $\sigma_k^2, k \leq n$ слабо коррелированы, то прошлые величины $h_{n-1}^2, h_{n-2}^2, \dots$ надо учитывать с малыми, убывающими весами. Если же величины $\sigma_k^2, k \leq n$ сильно коррелированы, то значения $h_{n-1}^2, h_{n-2}^2, \dots$ могут дать существенную дополнительную информацию (к той, что есть в h_n^2) о значениях σ_n^2 .

Эта идея приводит к рассмотрению экспоненциально-взвешенных оценок $\tilde{\sigma}_n^2 = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k h_{n-k}^2, 0 < \lambda < 1$, конструирование которых относит момент начала из нуля в $-\infty$. В этом предположении $\Delta_n = \frac{a_0}{1 - a_1} + c \sum_{k=0}^{\infty} a_1^k \delta_{n-k}$, где ряд сходится в среднем квадратическом. Заметим, что $(1 - \lambda) \sum_{k=0}^{\infty} \lambda^k = 1$, то есть сумма взвешенных коэффициентов при конструировании $\tilde{\sigma}_n^2$ равна единице.

Поскольку $E h_{n-k}^2 = E \sigma_{n-k}^2 = E \sigma_n^2$, то мы видим, что $E \tilde{\sigma}_n^2 = E \sigma_n^2$, то есть оценка $\tilde{\sigma}_n^2$ наряду с $\hat{\sigma}_n^2$ является несмещённой.

Отметим также, что точность оценки $\tilde{\sigma}_n^2$ сильно зависит от значения выбираемого параметра λ , поэтому возникает задача выбора оптимального значения параметра λ .

Значения $\tilde{\sigma}_n^2$ подчиняются рекуррентным соотношениям $\tilde{\sigma}_n^2 = \lambda \tilde{\sigma}_{n-1}^2 + (1 - \lambda) h_n^2$, которые удобны при отыскании оценок методами статистического анализа и моделирования.

Опираясь сейчас с моделью $h_n = e^{\frac{\Delta_n}{2}} \varepsilon_n$, где $\Delta_n = a_0 + a_1 \Delta_{n-1} + c \delta_n$, естественно было бы для оценки величин $\Delta_n = \ln \sigma_n^2$ воспользоваться оптимальной в среднеквадратическом смысле оценкой $m_n = E(\Delta_n / h_1, \dots, h_n)$.

К сожалению, нелинейность рассматриваемой схемы делает задачу отыскания m_n в явном виде очень сложной. Поэтому сначала необходимо линеаризовать рассматриваемую задачу и далее воспользоваться теорией гауссовской линейной фильтрации Калмана-Бьюси.

Имеем:

$$\ln h_n^2 = \ln \varepsilon_n^2 + \ln \sigma_n^2 = E \ln \varepsilon_n^2 + \Delta_n + (\ln \varepsilon_n^2 - E \ln \varepsilon_n^2), \text{ где } E \ln \varepsilon_n^2 \approx -1.27,$$

$D \ln \varepsilon_n^2 = \frac{\pi^2}{2} \approx 4.93$. Поэтому, обозначая $x_n = \ln^2 h_n$, получаем следующую линейную систему:

$$\Delta_n = a_0 + a_1 \Delta_{n-1} + c \delta_n, \quad x_n = -1.27 + \Delta_n + \frac{\pi}{\sqrt{2}} \xi_n, \text{ где } \xi_n = \frac{\sqrt{2}}{\pi} (\ln \varepsilon_n^2 - E \ln \varepsilon_n^2) \text{ с}$$

$$E\xi_n = 0, D\xi_n = 1.$$

Таким образом, мы получили линейную систему, в которой распределения величин ξ_n не являются гауссовскими, что не даёт возможности непосредственного применения теории линейной фильтрации Калмана-Бьюси.

Тем не менее, рассмотрим фильтр Калмана-Бьюси так, как если бы величины ξ_n были нормально распределёнными, $\xi_n \sim N(0,1), n \geq 1$, и независимыми от последовательности (δ_n) .

При этом допущении $\mu_n = E(\Delta_n / x_1, \dots, x_n)$ и $\gamma_n = D\Delta_n$. Тогда соответствующая система, определяющая эволюцию величин μ_n и γ_n , имеет следующий вид:

$$\mu_{n+1} = (a_0 + a_1\mu_n) + \frac{a_1\gamma_n}{\pi^2/2 + \gamma_n} (x_{n+1} + 1.27 - \mu_n)$$

$$\gamma_{n+1} = (a_1^2\gamma_n + c^2) - \frac{(a_1\gamma_n)^2}{\pi^2/2 + \gamma_n}$$

$$\text{При этом } \mu_0 = E\Delta_0, \gamma_0 = D\Delta_0.$$

Заметим, что если параметр c большой, то член Δ_n в x_n будет играть доминирующую роль, и в этом случае величины m_n хорошо аппроксимируются величинами μ_n .

Рассмотрим хаотические системы, то есть описываемые детерминистическими нелинейными системами, которые могут приводить к эффектам типа кластерности, наблюдаемому при статистическом анализе финансовых данных. Приведём некоторые примеры нелинейных хаотических систем, чтобы дать представление об их поведении, а также о том, как определить, порождается данная реализация стохастической или хаотической системой.

К сожалению, нелинейные хаотические модели плохо прогнозируемы. Причина в том, что, несмотря на детерминистический характер, поведение траекторий хаотических систем может сильно изменяться в зависимости от точности определения значений начальных данных и значений параметра λ .

Пример 1.

Рассмотрим так называемое логистическое отображение $x \rightarrow Tx \equiv \lambda x(1-x)$ и порождаемую им одномерную нелинейную динамическую систему $x_n = \lambda x_{n-1}(1-x_{n-1}), n \geq 1, 0 < x_0 < 1$. Впервые логистические уравнения появились в моделях популяционной динамики, учитывающих ограничения на рост численности популяций.

Для значений $\lambda \leq 1$ решения $x_n = x_n(\lambda) \rightarrow 0$ при $n \rightarrow \infty$ при всех $0 < x_0 < 1$. Таким образом, состояние $x_\infty = 0$ можно в этом случае рассматривать как то единственное устойчивое состояние, к которому сходятся значения x_n при $n \rightarrow \infty$.

При $\lambda = 2$ значения $x_n \rightarrow \frac{1}{2}$. Следовательно, в этом случае также существует единственное устойчивое состояние $x_\infty = 1/2$, к которому притягиваются значения x_n при $n \rightarrow \infty$.

Будем теперь увеличивать значение параметра λ . При $\lambda < 3$ у системы всё ещё будет только одно устойчивое состояние. Однако, при $\lambda = 3$ возникает качественно новый эффект – по мере увеличения n появляются два состояния устойчивости x_∞ , в которых попеременно находится система.

Такой же характер поведения у системы будет сохраняться и при увеличении значений параметра λ . Но при $\lambda = 3.4494\dots$ с системой начинает происходить нечто новое – у неё появляются четыре состояния устойчивости, по которым происходит движение системы.

При увеличении λ у системы начинают появляться новые состояния устойчивости: при $\lambda = 3.5644\dots$ этих состояний 16, при $\lambda = 3.5696\dots$ их уже 64. А при $\lambda = 3.6$ число таких состояний становится равным бесконечности, что интерпретируют как потерю системой устойчивости и переход системы в состояние хаоса.

При этом полностью исчезает периодический характер смены состояний, и система начинает совершать блуждание по бесконечному числу состояний, прыгая из одного из них в другое. Отметим, что, хотя система и является детерминистической, практически невозможно предсказать, где окажется система через некоторое время, поскольку ограниченная точность в определении значений x_n и λ может сильно повлиять на значения прогнозируемых величин.

Значения (λ_k) параметра λ , где происходит разветвление (бифуркация), становятся всё ближе и ближе.

М. Фейгенбаум высказал гипотезу и О. Ланфорд доказал, что для всех параболических систем $\frac{\lambda_k - \lambda_{k-1}}{\lambda_{k+1} - \lambda_k} \rightarrow F, k \rightarrow \infty$, где $F = 4.669201\dots$ – универсальная константа (число Фейгенбаума).

Параметр $\lambda = 4$ играет особую роль. Именно при этом значении последовательность наблюдений соответствующей хаотической последовательности (x_n) напоминает реализацию стохастической последовательности типа белого шума. В самом деле, возьмём $x_0 = 1$ и подсчитаем рекуррентным образом $x_1, x_2, \dots, x_{1000}$. Подсчитанные по этим 1000 значениям эмпирические среднее значение и стандартное отклонение дают, соответственно, значения 0.48887 и 0.35742 (с точностью до пяти знаков). В следующей таблице приведены значения эмпирической корреляционной функции $\hat{\rho}(k)$, подсчитанной по значениям $x_1, x_2, \dots, x_{1000}$. Из этой таблицы видно, что величины (x_n) , порождённые логистическим отображением с $\lambda = 4$, практически можно считать некоррелированными, и в этом смысле последовательность (x_n) может быть названа хаотическим белым шумом.

№	Значение
1	-0.033
2	-0.058
3	-0.025
4	-0.035
5	-0.012
6	-0.032
7	-0.048
8	0.027
9	-0.020
10	-0.013
11	-0.046
12	0.002
13	-0.011
14	0.040
15	0.014
16	-0.023
17	-0.030
18	0.037
19	0.078
20	0.017
21	-0.008
22	0.009
23	-0.039
24	-0.020
25	-0.008
26	0.017
27	0.006
28	-0.004
29	-0.019
30	-0.076
31	0.038
32	-0.017
33	0.014
34	0.001
35	0.017
36	-0.052
37	0.004
38	0.053
39	-0.021
40	0.007

Таблица 1.

Интересно отметить, что для системы $x_n = 4x_{n-1}(1-x_{n-1}), n \geq 1$ с $x_0 \in (0,1)$ существует инвариантное распределение P (то есть такое, что $P(T^{-1}A) = P(A)$ для любого борелевского множества A из $(0,1)$), плотность которого

$$p(x) = \frac{1}{\pi[x(1-x)]^{1/2}}, x \in (0,1).$$

Тем самым, если считать начальное значение x_0 случайной величиной с плотностью распределения вероятностей $p = p(x)$, то случайные величины $x_n, n \geq 1$ будут иметь то же самое распределение, что и x_0 . Полезно подчеркнуть, что у получаемой таким образом стохастической динамической системы (x_n) вся случайность полностью определяется случайным начальным значением x_0 , а динамика переходов $x_n \rightarrow x_{n+1}$ задаётся детерминированным образом.

Нетрудно найти, что $Ex_0 = \frac{1}{2}, Ex_0^2 = \frac{3}{8}, Dx_0 = \frac{1}{8} = (0.35355\dots)^2$ (сравните со значениями 0.48887 и 0.35742) и $\rho(k) = \frac{Ex_0 x_k - Ex_0 Ex_k}{\sqrt{Dx_0 Dx_k}} = \begin{cases} 1, k = 0 \\ 0, k \neq 0 \end{cases}$.

Пример 2. Преобразование Бернулли.

$$x_n = 2x_{n-1} \pmod{1}, x_0 \in (0,1)$$

Инвариантным здесь является равномерное распределение с плотностью $p(x) = 1, x \in (0,1)$. При этом $Ex_0 = 1/2, Ex_0^2 = 1/3, Dx_0 = 1/12, \rho(k) = 2^{-k}, k = 0, 1, \dots$

Пример 3. Палаточное преобразование.

$$x_n = 1 - |1 - 2x_{n-1}|, x_0 \in (0,1)$$

Как и в примере 2, инвариантным здесь является равномерное распределение на $(0,1)$. При этом $Ex_0 = 1/2, Ex_0^2 = 1/3, Dx_0 = 1/12, \rho(k) = 0, k \neq 0$.

Пример 4.

$$x_n = 1 - 2\sqrt{|x_{n-1}|}, x_0 \in (-1,1)$$

Инвариантным здесь является распределение на $(-1,1)$ с плотностью $p(x) = (1-x)/2$. При этом $Ex_0 = -1/3, Ex_0^2 = 1/3, Dx_0 = 2/9$.

На примере логистической системы, развивающейся бинарным образом, чётко прослеживается идея фрактальности. Кроме этого, поведение хаотических систем наводит на мысль об их использовании при построении моделей эволюции финансовых индексов, особенно в кризисные периоды, которым присуща скорее именно хаотичность, а не стохастичность.

Рассмотрим проблематику различимости хаотических и стохастических последовательностей. Известно, что чисто детерминистические динамические системы могут проявлять свойства стохастического белого шума. Поэтому зададимся вопросом, как различать стохастические и хаотические последовательности.

Рассмотрим функцию $C(\varepsilon) = \lim_{N \rightarrow \infty} \frac{\psi(N, \varepsilon)}{N^2}$, где $\psi(N, \varepsilon)$ – число тех пар $(i, j), i, j \leq N$, для которых в рассматриваемой последовательности (x_n) $|x_i - x_j| < \varepsilon$.

Помимо функции $C(\varepsilon)$, привлекаются к рассмотрению также и функции $C_m(\varepsilon) = \lim_{N \rightarrow \infty} \frac{\psi_m(N, \varepsilon)}{N^2}$, где $\psi_m(N, \varepsilon)$ – число тех пар (i, j) , для которых все компоненты векторов $(x_i, x_{i+1}, \dots, x_{i+m-1})$ и $(x_j, x_{j+1}, \dots, x_{j+m-1})$ с $i, j \leq N$ отличаются не более чем на ε . В случае $m = 1$ имеем $\psi_1(N, \varepsilon) = \psi(N, \varepsilon)$.

Для стохастических последовательностей (x_n) типа белого шума при малых ε функция $C_m(\varepsilon) \sim \varepsilon^{\nu_m}$, где фрактальный показатель $\nu_m = m$. Этим свойством обладают также и многие детерминистические системы. Показатель ν_m носит также название корреляционной размерности и тесно связан с хаусдорфовой и информационной колмогоровской размерностями.

Идея различения хаотических и стохастических последовательностей основана на том наблюдении, что корреляционная размерность у таких последовательностей различна. А именно, у стохастических последовательностей она больше, чем у хаотических.

В качестве оценок корреляционной размерности ν_m естественно взять величины $\tilde{\nu}_{m,j} = \frac{\ln C_m(\varepsilon_j) - \ln C_m(\varepsilon_{j+1})}{\ln \varepsilon_j - \ln \varepsilon_{j+1}}$ или $\tilde{\nu}_m(k) = \frac{1}{k} \sum_{j=1}^k \tilde{\nu}_{m,j}$, где $\varepsilon_j = \varphi^j, 0 < \varphi < 1$.

Следующая таблица даёт для логистической последовательности $(x_n)_{n \leq N}$ с $N = 5900$ значения $\tilde{\nu}_{m,j}$ для $m = 1, 2, 3, 4, 5, 10$ и $\varepsilon_j = \varphi^j$ с $\varphi = 0.9$ для ряда значений j .

Сравним результаты этой таблицы с теми оценками для $\tilde{\nu}_{m,j}$, которые получаются при моделировании гауссовского белого шума с теми же параметрами, что и для логистического отображения.

Из сопоставления этих таблиц видно, что в случае $m = 1$ весьма трудно по значениям корреляционной размерности $\tilde{\nu}_{1,j}$ отличить хаотичность и стохастичность. Привлечение же больших значений m показывает существенное различие в величинах $\tilde{\nu}_{m,j}$ для хаотического и стохастического случаев, что может служить достаточно весомым основанием для утверждения о разной природе образования соответствующих последовательностей (x_n) , хотя с точки зрения их эмпирических средних, дисперсий и корреляций разницы практически нет.

Для иллюстрации проблематики различимости стохастичности и хаотичности в финансовых рядах приведём таблицы значений корреляционных размерностей для дневных величин возврата для индексов IBM и S&P500 по 5903 наблюдениям в период времени 2.07.1962-31.12.1985 г.г. Из сопоставления этих двух таблиц видна однородность фрактальной структуры.

Таблица 2.

Значения $\tilde{V}_{m,j}$ для
логистической системы.

m, j	1	2	3	4	5	10
20	0.78	0.90	0.96	0.98	1.02	1.19
30	0.81	0.89	0.95	0.98	0.98	1.11
35	0.83	0.90	0.94	0.97	0.95	1.01
40	0.83	0.91	0.97	0.99	1.05	1.20

Таблица 3.

Значения $\tilde{V}_{m,j}$ для
гауссовского белого шума.

m, j	1	2	3	4	5	10
20	0.84	1.68	2.52	3.35	4.20	8.43
30	0.98	1.97	2.95	3.98	4.98	-
35	0.99	1.97	2.93	4.00	5.53	-
40	1.00	2.02	3.03	4.15	5.38	-

Таблица 4.

Значения $\tilde{V}_{m,j}$ для ИВМ.

m, j	1	2	3	4	5	10
20	0.46	0.90	1.31	1.68	2.05	3.63
30	0.83	1.76	2.61	3.44	4.27	8.44
35	0.97	1.93	2.88	3.82	4.79	9.84
40	0.98	1.96	2.94	3.86	4.94	-

Таблица 5.

Значения $\tilde{V}_{m,j}$ для S&P500.

m, j	1	2	3	4	5	10
20	0.58	1.10	1.58	2.03	2.43	3.93
30	0.93	1.82	2.07	3.49	4.25	6.93
35	0.98	1.94	2.88	3.79	4.75	11.00
40	0.99	1.98	2.92	3.84	4.81	-

В таблицах 2,3,4,5 отражены значения оценок для различных индексов.

Сравнение таблиц 4 и 5 с таблицами 2 и 3 показывает, что для индексов IBM и S&P500 последовательности (h_n) с $h_n = \ln \frac{S_n}{S_{n-1}}, n \geq 1$ ведут себя скорее

как стохастический белый шум, хотя это не отвергает гипотезу о том, что близкими свойствами могут обладать и другие хаотические последовательности с большей корреляционной размерностью.

Рассмотрим ещё один подход для выявления различий между стохастичностью и хаотичностью.

Пусть $x = (x_n)$ – хаотическая последовательность, порождённая некоторой динамической системой с распределением вероятностей $F = F(x)$ для x_0 , являющимся инвариантным для данной системы.

Пусть теперь $\tilde{x} = (\tilde{x}_n)$ – стохастическая последовательность, состоящая из независимых одинаково распределённых величин с одномерным распределением $F = F(x)$.

Образует величины $M_n = \max(x_0, x_1, \dots, x_n)$ и $\tilde{M}_n = \max(\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_n)$ и пусть $F_n(x) = P(M_n \leq x), \tilde{F}_n(x) = P(\tilde{M}_n \leq x)$.

Идея подхода основана на том наблюдении, что максимум служит хорошей характеристикой, позволяющей улавливать отличия стохастических последовательностей от хаотических.

В теории предельных теорем для экстремальных значений хорошо известны необходимые и достаточные условия, при которых величины $a_n(\tilde{M}_n - b_n), n \geq 1$ имеют нетривиальное предельное распределение $\lim_n P(a_n(\tilde{M}_n - b_n) \leq x) = \tilde{G}(x)$ для некоторых констант $a_n > 0$ и $b_n, n \geq 1$.

Имея распределение $\tilde{F}_n(x) = (F(x))^n$ и предельное распределение $\tilde{G}(x)$, естественно было бы сравнить их с соответствующими распределениями $F_n(x)$ и, если это возможно, с их пределами $G(x)$.

Здесь имеется большая техническая трудность, связанная с тем, что для рассмотренных примеров отсутствуют удобные для анализа аналитические выражения для $F_n(x)$. Ввиду этого обстоятельства, мы пойдём по пути численного анализа распределений $F_n(x)$ для больших n и сравнения их с соответствующими распределениями $\tilde{F}_n(x) = |F(x)|^n$.

Для динамических систем этот анализ показывает, что в глобальном смысле поведение $F_n(x)$ для хаотических систем с инвариантным распределением $F(x)$ качественно отличается от поведения $\tilde{F}_n(x)$ для стохастических систем, образованных независимыми одинаково распределёнными величинами с одномерным распределением $F(x)$. Это говорит о том, что для рассматриваемых моделей максимум является хорошей статистикой в рассматриваемой проблеме различимости хаотичности и стохастичности. Но это не исключает того, что может найтись хаотическая система, которую будет трудно отличить от стохастического белого шума.