Лекция 1

Введение в теорию временных рядов. Случайные процессы. Стационарные случайные процессы.

Данные и модель

Термин «временной ряд» используется для обозначения двух объектов:
1. Совокупность данных (наблюдений), собранных в определённый момент времени *t*: $\{x_t \mid t \in T\}$, где T — индексное множество. Обычно $T = [a; b]$ для непрерывных наблюдений и $T = \{t_1, t_2, \dots\}$ для дискретных наблюдений. В этой лекции мы рассмотрим случай $T \subset \mathbb{Z}$.

2. Модель, описывающая данные. Для этого мы будем использовать случайный процесс (см. определение ниже) $\{X_t \mid t \in T\}$. Общая идея: каждый x_t является реализацией случайной величины X_t.

Обратите внимание, что для выборки данных мы обычно имеем $\{x_1, \dots, x_n\}$ (мы начинаем наблюдение в момент времени 1 и останавливаем в момент времени n), в то время как в качестве модели мы используем $\{X_t \mid t \in \mathbb{Z}\}$, т.е. процесс происходит из бесконечного прошлого в бесконечное будущее. Интерпретация такова: процесс обычно уже запущен до того, как мы начнём наблюдать за выходными данными (до момента времени 1), и он продолжится после того, как мы прекратим наблюдение в момент времени n. Техническая причина: математическое удобство.

Чем анализ временных рядов отличается от классической статистики?

Данные зависимы и коррелируют, тогда как в статистике обычно предполагается независимость между наблюдениями. Грубо говоря, в статистике экспериментатор может повторить эксперимент в тех же условиях и независимо. Во временных рядах мы не можем повторить эксперимент, мы только наблюдаем за запущенным процессом. С одной стороны, это требует более сложных моделей, но с другой стороны, зависимость является важнейшим компонентом прогнозирования, одной из центральных тем анализа временных рядов.

Определение. Случайный процесс — это совокупность случайных величин $\{X_t \mid t \in T\}$, определённых на общем вероятностном пространстве $(\Omega, \mathscr{F}, P)$. Функции $\{X_t(\omega) \mid t \in T\}$, где $\omega \in \Omega$, называются **реализациями, траекториями** или также **траекториями выборки**.
Мы используем $\{x_t \mid t \in T\}$ для обозначения наблюдаемых значений.

Обозначения временного ряда как X_t и $X(t)$ взаимозаменяемы.

Определение. (**Белый шум**). Процесс $\{Z_t, t \in \mathbb{Z}\}$ называется **белым шумом**, если выполняются следующие условия:

1. $Var(Z_t) = \sigma_Z^2 < \infty, \qquad \forall t,$
2. $EZ_t = \mu, \quad \forall t,$
3. $Cov(Z_t, Z_{t+h}) = 0 \quad \forall t, |h| \geq 1.$

Белый шум часто обозначается : $Z_t \sim WN(\mu, \sigma^2)$

Обратите внимание, что белый шум может быть (i) зависимым! Например, известная модель GARCH, используемая в эконометрических приложениях, представляет собой всего лишь белый шум. (ii) Распределение маргинальных значений может быть совершенно разным для разных моментов времени.

Определение. (**Конечномерные распределения**). Пусть $\{X_t, t \in T\}$ — случайный процесс, и пусть

$$\mathscr{T} = \{\mathbf{t} = (t_1, \ldots, t_n),\ t_1 < t_2 < \ldots < t_n, \quad n = 1,2,...\}.$$

Тогда **конечномерными функциями распределения случайного процесса** являются функции $\{F_t(\,\cdot\,),\ t = (t_1, \ldots, t_n) \in \mathscr{T}\}$, где

$$F_t(x) = F_t(x_1, \ldots, x_n) = P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n).$$

Пример . Если случайные величины X_t независимы, то конечномерные функции распределения имеют вид

$$F_t(\mathbf{x}) = F_t(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_{t_i}}(x_i) = P(X_{t_1} \leq x_1) \cdots P(X_{t_n} \leq x_n)$$

Теорема Колмогорова утверждает, что если заданы свойства конечномерных распределений (согласованность), то такие распределения действительно определяют случайный процесс на некотором вероятностном пространстве.

Теорема Колмогорова о согласованных конечномерных распределениях случайного процесса
(Kolmogorov consistensy theorem)

$$F_{t_1 \ldots t_k}(y_1, \ldots, y_k) = \lim_{y_{k+1} \to \infty} F_{t_1 \ldots t_k, t_{k+1}}(y_1, \ldots, y_k, y_{k+1})$$

Стационарность

В контексте конечного числа случайных величин мы часто используем ковариационную функцию, чтобы понять зависимость между ними. Теперь нам нужен инструмент, который может расширить концепцию ковариационной матрицы для работы с бесконечным числом случайных величин. Это автоковариационная функция.

Определение (Автоковариационная функция). Для процесса $\{X_t, t \in T\}$, для которого Var (Xt) < ∞ при каждом t ∈ T, автоковариационная функция γX(··) процесса {Xt} задаётся формулой:

$$\gamma_X(r, s) = Cov(X_r, X_s) = E[(X_r - E[X_r])(X_s - E[X_s])], \quad \forall r, s \in T.$$

Замечание: Обратите внимание, что $Var(X) < +\infty \iff EX^2 < +\infty$.
Последнее означает (согласно неравенству Коши-Шварца, предложение 2.1.1), что $E|X| < +\infty$.

Привлекательным аспектом независимых одинаково распределенных последовательностей является то, что их стохастическое поведение остаётся стабильным. Теперь мы вводим схему, которая очень часто используется для описания определенной структурной устойчивости с точки зрения первых двух моментов процесса и ее зависимости.

Теория временных рядов основана на предположении о **стационарности в широком смысле**, которую в литературе также называют **слабой стационарностью**, или **стационарность второго порядка.**

Определение (Слабая стационарность). Временной ряд $\{X_t, t \in \mathbb{Z}\}$называется **стационарным,** если:
(i) $EX_t^2 < +\infty, \forall t$
(ii) $EX_t = m, \ \forall t \in \mathbb{Z}$ (т.е. Математическое ожидаемое значение постоянно)
(iii) $\gamma_X(r, s) = Cov(X_r, X_s) = \gamma_X(r + t, s + t), \quad \forall t \in \mathbb{Z}, \quad \forall r, s \in \mathbb{Z}$.

! **В дальнейшем слабо стационарный временной ряд будем называть стационарным.**

Реальные данные часто не являются стационарными: например, они демонстрируют линейный тренд во времени или имеют сезонный эффект.
Пример. Последовательность iid слабо стационарна (при наличии вторых моментов), тогда как модель случайного блуждания не является слабо стационарной.

Замечание. Очевидно, $\gamma_X(r,s) = \gamma_X(s,r)$. Следовательно, согласно (iii)

$$\gamma_X(0,s-r) = \gamma_X(r,s) = \gamma_X(r-s,0) = \gamma_X(0,r-s)$$

Это означает, что ACF стационарного временного ряда может быть выражена как однопараметрическая функция $\gamma_X(h)$, где $h = |r-s|$. В этом случае для оценки зависимости часто используют автокорреляционную функцию (ACF):

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{Cov(X_{t+h}, X_t)}{Var(X)} = Corr(X_{t+h}, X_t)$$

Замечание. $\rho_X(h)$ — хорошая и простая мера зависимости.
Но $\rho_X(h)$= 0 обычно не означает, что X_t не зависит от X_{t+h}.
Для гауссовских процессов некоррелированность подразумевает независимость.
Вот ещё одно, более строгое понятие стационарности.

Определение. (Строгая стационарность). Процесс $\{X_t, t \in \mathbb{Z}\}$ является строго стационарным, если
$F_t(\cdot) = F_{t+h}(\cdot), \quad \forall h \in Z, \quad t \in \mathcal{T}$ или, что эквивалентно,

$$P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n) = P(X_{t1+h} \leq x_1, \ldots, X_{tn+h} \leq x_n).$$

Другими словами, Процесс $\{X_t, t \in \mathbb{Z}\}$ является **строго стационарным** или **стационарным в узком смысле,** или имеет **стационарность первого порядка**, если $(X_{t_1}, \ldots, X_{t_k})$ и $(X_{t_1+h}, \ldots, X_{t_k+h})$ имеют одинаковое распределение для всех наборов временных точек $t_1, \ldots, t_k$ и всех целых чисел h.

Пример. Последовательность независимых случайных чисел строго стационарна.
Пример . $X_t = X, \quad \forall t$ строго стационарна.
Следующее утверждение весьма полезно:

Предложение. Предположим, что $\{X_t, t \in \mathbb{Z}\}$ строго стационарна и $f : \mathbb{R}^\infty \longrightarrow \mathbb{R}$ является измеримой по Борелю функцией. Тогда $Y_t = f(X_t, X_{t-1}, \ldots)$ является строго стационарной.

Пример . Предположим, что $\{\varepsilon_t\} \sim IID$. Тогда это строго стационарный процесс. Определим линейный процесс
$X_t = \sum_{k \geq 0} a_k \varepsilon_{t-k} = a_0 \varepsilon_t + a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2} + \cdots = f(\varepsilon_t, \varepsilon_{t-1}, \ldots)$
Если ряд сходится (в некотором смысле, который будет указан), то $\{X_t, t \in \mathbb{Z}\}$ строго стационарен.

Связь между слабой и строгой стационарностью. Если $\{X_t\}$ строго стационарна, то распределение F_t для X_t одинаково для каждого $t \in \mathbb{Z}$. Более того, любая пара (X_t, X_{t+h}) имеет совместное распределение $F_{t,t+h}$, которое не зависит от t. Следовательно,

$$EX_t = \int_{\mathbb{R}} x dF_t(x) = m$$

не зависит от t и

$$Cov(X_t, X_{t+h}) = \int_{\mathbb{R}} \int_{\mathbb{R}} (x-m)(y-m) dF_{t,t+h}(x,y)$$

Следовательно:
Строгая стационарность $\Longrightarrow$ слабая стационарность при условии, что последовательность имеет конечные вторые моменты.

Обратное в общем случае неверно. (Найдите простой контрпример.)

Существует важный случай, когда слабая стационарность подразумевает строгую стационарность. Это **случай гауссовского процесса.** В этом случае слабая стационарность $\{X_t\}$ подразумевает строгую стационарность, поскольку $(X_{t_1}, \ldots, X_{t_n})'$ и $(X_{t_1+h}, \ldots, X_{t_n+h})'$ имеют одинаковые средние значения и ковариационную матрицу, а следовательно, и одинаковое распределение для всех $n = \{1,2,...\}$ и для всех $h, t_1, \ldots, t_n \in \mathbb{Z}$.

ПРИМЕЧАНИЕ: **В дальнейшем мы будем называть временной ряд стационарным, если он слабо стационарен.**

Свойства автоковариационной функции

Утверждение. (Элементарные свойства автоковариационной функции).

Если $\gamma_X(\cdot)$ — автоковариационной функция стационарного процесса $\{X_t, t \in \mathbb{Z}\}$, то
(i) $\gamma_X(0) \geq 0$,
(ii) $|\gamma_X(h)| \leq \gamma_X(0), \quad \forall h \in \mathbb{Z}$
(iii) $\gamma_X(h) = \gamma_X(-h), \quad \forall h \in \mathbb{Z}$.

Доказательство. (i) следует из $\gamma(0) = Var(X_t) \geq 0$,
(ii) следует из
$|\gamma_X(h)| = |Cov(X_{t+h}, X_t)| \leq (Var(X_{t+h}))^{\frac{1}{2}}(Var(X_t))^{\frac{1}{2}} = \gamma_X(0)$
(неравенство Коши-Буняковского (Шварца), и
(iii) тривиально

$\gamma_X(h) = Cov(X_{t+h}, X_t) = Cov(X_t, X_{t-h}) = \gamma_X(-h), \quad \forall h \in \mathbb{Z}$.

Определение. (**Неотрицательная определённость, н.о**.).
Действительнозначная функция от целых чисел $\kappa : \mathbb{Z} \to \mathbb{R}$ называется неотрицательно определённой (н.о.), если

$$\sum_{i=1}^{n} \sum_{i=1}^{n} a_i \cdot \kappa(t_i - t_j) \cdot a_j \geq 0, \quad a_i \in \mathbb{R},\ t_i \in \mathbb{Z}, \quad n \geq 1$$

Другими словами: матрицы $K = ((\kappa(t_i - t_j)))_{i,j=1}^{n}$i,j=1 являются н.о.
Теорема. (Характеристика ACF). Функция $\gamma : \mathbb{Z} \longrightarrow \mathbb{R}$ является автокорреляционной функцией стационарного процесса $\{X_t, t \in \mathbb{Z}\}$, тогда и только тогда, когда γ чётна и н.о.
Доказательство. Предположим, что γ — ACovF стационарного процесса. Тогда γ чётна (Утверждение , (iii)). Далее

$$0 \leq Var\left(\sum_{i=1}^{n} a_i X_{t_i}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j Cov(X_{t_i}, X_{t_j}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(t_i - t_j)$$

Предположим, что γ чётно и н.о(неотрицательно определенная) . Покажем, что существует гауссовский процесс с ACovF γ. Определим $\Gamma_t = (\gamma(t_i - t_j))_{i,j}^{n}$. Заметим, что, поскольку Γ_t нечётно, мы можем найти многомерный нормальный вектор X с $\Gamma_t = Cov(\mathbf{X})$. Теперь воспользуемся « Фурье версией» теоремы Колмогорова о продолжении: пусть $\varphi_t(\mathbf{u})$ — характеристическая функция нормально распределенного $N(\mathbf{0}, \Gamma_t)$ случайного вектора .