ЛЕКЦИЯ 1

Случайные процессы: определения

В этой лекции мы вводим понятия случайного процесса, его основных характеристик (конечномерные распределения, матожидание, ковариационная функция), стационарности в узком и широком смысле, а также обсуждаем первые примеры.

1.1 Случайный процесс

В мире достаточно большое число динамических процессов, которые имеют в себе некоторую неопределённость.

1. Датчики ЭКГ(электрокардиография), как и любые датчики выдают на выходе зашумлённые данные. Замеренное датчиком напряжение в момент времени $t \in \mathbb{R}_+$ – это X_t, случайная величина со значением в $\mathbb{R}$.

2. Очереди на кассах в супермаркете случайные и меняются со временем. Мы не можем знать, в какой момент к нам придут клиенты. Число клиентов X_t в момент времени $t \in \mathbb{R}_+$ – это случайная величина принимающая целые, но сколь угодно большие значения, ограниченные только размером магазина и прилегающей территории.

3. Сам клиент (с позиции руководителя магазина) – это тоже очень неопределённый субъект. Мы знаем, что он может случайно переходить по отделам, но достоверно не известно, как именно. При этом очевидно, что магазин может влиять на модель клиента путём разных управленческих решений (правильного размещения отделов, скидок и пр.). Для выручки магазина может быть не важно, сколько времени клиент провёл на каждом отделе, но важно, как именно он между ними перемещался. Таким образом, положение клиента X_t в момент $t \in \mathbb{Z}_+$ – это случайная величина, принимающая значения в конечном множестве $\{1, .., n\}$, если в магазине всего n отделов.

4. Давление в газопроводе зависит от расстояния от точки замера до источника газа очень неопределённым образом в силу засорённости трубы и её геометрии. Давление X_t в точке трубы $t \in D \subset \mathbb{R}^4$ – это случайная величина, принимающая значения в $\mathbb{R}_+$.

Очевидно, что одних случайных величин становится недостаточно для описания динамики таких систем – например, положение клиента в магазине зависит от того, где он был раньше и отчасти может определять, где он будет потом. Поэтому на все X_t нужен некоторый *общий* взгляд. Для начала заметим, что если мы зафиксируем время (или точку пространства, как в примере 4), то показатели (напряжение, размер очереди на кассе, давление газа, номер отдела, где находится клиент) – это случайные величины, которых очень много и при этом они ещё нетривиально зависят друг от друга.

Определение 1.1. *Пусть* $(\Omega, \mathcal{F}, P)$ *– вероятностное пространство. Семейство случайных величин* $X_t : \Omega \to \Xi$, *принимающих значения в измеримом пространстве* $(\Xi, \mathcal{G})$, *проиндексированное значениями* $t \in T$ *из некоторого множества* T, *называют случайным процессом и обозначают* $(X_t)_{t\in T}$. *Часто для краткости мы будем писать* X.

В примерах 1 и 2 индексное пространство $T = \mathbb{R}_+$, в третьем примере $T = \mathbb{Z}_+$; так, индекс – это время. Тем не менее, часто индексом может быть, например, подмножество $\mathbb{R}^3$, как в случае с примером 4. Пространство значений $\Xi = \mathbb{R}_+$ в примерах 1 и 4, в примере 2 это $\Xi = \mathbb{Z}$, а в третьем примере $\Xi = \{1, .., n\}$.

Случайный процесс обозначают как $X_t(\omega)$ или $X(t, \omega)$, таким образом явно обозначая, что его можно рассматривать как функцию $X : T \times \Omega \to \Xi$. Если мы фиксируем индекс $t = t_0$, то мы получаем $X_{t_0} : \Omega \to \mathbb{R}$, случайную величину, называемую иногда *сечением процесса в точке* t_0 . Если же мы зафиксируем $\omega = \omega_0$, то $X_t(\omega_0) : T \to \mathbb{R}$ – это детерминированная функция индекса t, называемая *реализацией* или *траекторией* случайного процесса X. Иными словами, выпало событие $\omega \in \Omega$ и сразу реализовалась вся траектория процесса $X_t(w)$.

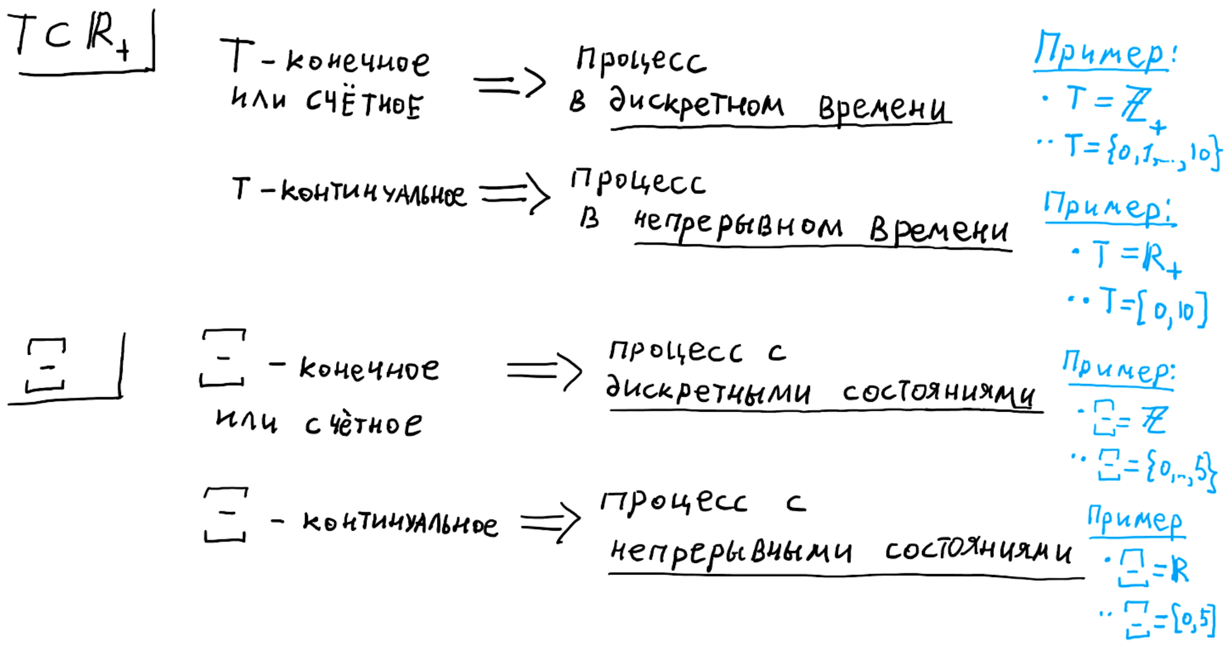

Рис. 1.1: Одна из известных классификаций случайных процессов

Для многих результатов существенно, как именно устроено индексное пространство T и множество значений Ξ. Рассмотрим одну известную классификацию (см. Рис. 1.1).

Пусть $T \subset \mathbb{R}_+$. Если T счётно (например, $T = \mathbb{Z}_+$), то говорят о процессе в *дискретном времени*; если T континуально (к примеру, $T = \mathbb{R}_+$), то X – процесс в *непрерывном времени*. Похожее разделение есть для пространства значений: если Ξ не более, чем счётно, то X – процесс с *дискретными состояниями*; если Ξ континуально, то X – процесс с *непрерывными состояниями*.

Пример 1.1. *Пусть $U \in \mathcal{U}[0,1]$, а случайный процесс задаётся $X_t = U sin(t)$. Траекториями такого процесса будут синусы со случайными амплитудами U (См. Рис. 1.4), а сечениями будут случайные величины $X_t \sim \mathcal{U}[0, \sin t]$ для $t \in (2\pi k, \pi + 2\pi k)$, $X_t \sim \mathcal{U}(\sin t, 0)$ для $t \in (\pi + 2\pi k, 2\pi + 2\pi k)$ и $X_t = 0$ для $t = 2\pi k$.*

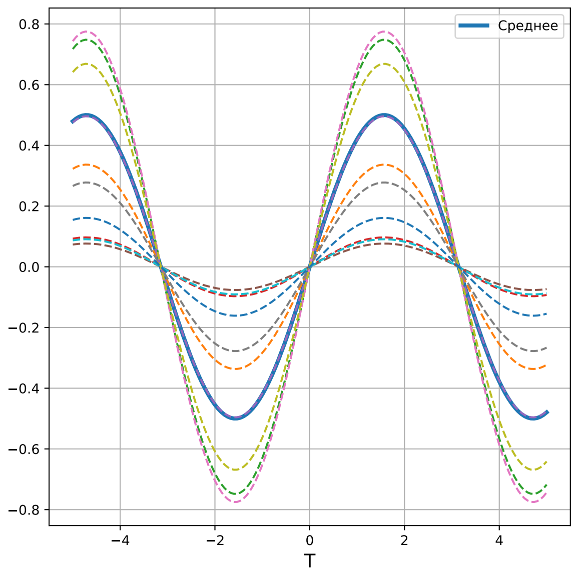

Рис. 1.2: Траектории процесса из Примера 1.1

1.2 Семейство конечномерных распределений и согласованность

Пусть $(\Omega, \mathcal{F}, P)$ – вероятностное пространство и $(X_t)_{t\in T}$ – случайный процесс на нём со значениями в измеримом пространстве $(\Xi, \mathcal{G})$. Если $\widetilde{\Omega}$ – пространство всех функций $\widetilde{\omega} : T \to \Xi$, то мы можем определить *цилиндрическое множество*, как

$$B_{t_1,..,t_n}^{A_1,..,A_n} := \left\{ \widetilde{\omega} \in \widetilde{\Omega} \ : \ \widetilde{\omega}(t_1) \in A_1, .., \widetilde{\omega}(t_n) \in A_n \right\} \tag{1.1}$$

для $t_1, .., t_n \in T$ и $A_1, .., A_n \in \mathcal{G}$. Если мы определим *сигма-алгебру цилиндров* $\mathcal{B}$ как порождённую всеми возможными цилиндрическими множествами, то мы получим, что каждый цилиндр – это измеримое множество. Более того, мы можем определить меру каждого цилиндра как

$$\widetilde{\mathrm{P}}\left(B_{t_1,..,t_n}^{A_1,..,A_n}\right) = P\left(\{\omega \in \Omega \;:\; X_{t_1}(\omega) \in A_1, .., X_{t_n}(\omega) \in A_n\}\right). \tag{1.2}$$

Таким образом, мы получили, что на основе случайного процесса X можно (с использованием теоремы о продолжении меры) задать вероятностное пространство $(\widetilde{\Omega}, \mathcal{B}, \widetilde{\mathrm{P}})$ на множестве траекторий.

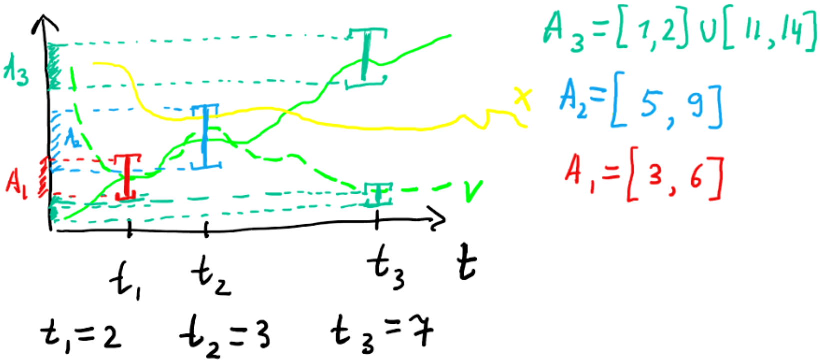

Рис. 1.3: Пример цилиндрического множества: две зелёных функции принадлежат цилиндру $B_{t_1,t_2,t_3}^{A_1,A_2,A_3}$, а жёлтая – нет

Поскольку в каждый фиксированный момент X_t – случайная величина, мы можем исследовать процесс и сложные зависимости между разными величинами $X_{t_1}, .., X_{t_n}$ изучая совместные распределения, которые имеют самое прямое отношение к пространству $(\widetilde{\Omega}, \mathcal{B}, \widetilde{\mathrm{P}})$.

Определение 1.2. *Пусть $(X_t)_{t\in T}$ – случайный процесс, определённый на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и принимающий значения в измеримом пространстве $(\Xi, \mathcal{G})$ и $t_1, .., t_n \in T$. Совместный закон распределения*

$$F(t_1, .., t_n; A_1, .., A_n) = F_{t_1,..,t_n}(A_1, .., A_n) := P\left(X_{t_1} \in A_1, .., X_{t_n} \in A_n\right)$$

называют n-мерным распределением в точках $t_1, .., t_n \in T$. Часто вводят также функцию распределения

$$F_{t_1,..,t_n}(x_1, .., x_n) := P\left(X_{t_1} \le x_1, .., X_{t_n} \le x_n\right).$$

Одномерных распределений оказывается недостаточно для описания случайного процесса – они ничего не говорят о зависимости случайных величин X_t в разных точках $t \in T$. Но в некоторых простых случаях их может быть достаточно чтобы задать полное семейство конечномерных распределений.

Упражнение 1.1. *Выпишите конечномерные распределения для процесса из Примера 1.1.*

Пример 1.2. *Случайный процесс X_t состоящий из независимых в совокупности (то есть, любые k сечений $X_{t_1},..,X_{t_k}$ независимы в совокупности) и одинаково распределённых случайных величин называется белым шумом. Например, если $X_t \sim \mathcal{N}(0,1)$ для всех $t \in T$ и любые сечения $X_{t_1},..,X_{t_k}$ независимы в совокупности, получим (стандартный) гауссовский белый шум.*

Так мы задали процесс X, задав семейство всех конечномерных распределений. Одномерные распределения $F(t_1;A_1) = f(A_1)$ фиксированы, а любое n-мерное распределение будет иметь вид

$$F(t_1,..,t_n;A_1,..,A_n) = \prod_{k=1}^{n} P(X_{t_k} \in A_k) = \prod_{k=1}^{n} f(A_k)$$

в силу условия независимости. В зависимости от приложений под белым шумом могут понимать разные предположения: иногда требуют дополнительно гауссовость распределения, а иногда вместо независимости только некоррелированность.

Пример 1.3. *Рассмотрим модель авторегрессии порядка 1 (обозначают $AR(1)$) с параметром $\alpha, x_0 \in \mathbb{R}$. Это процесс в дискретном времени. Пусть $T = \mathbb{Z}_+$, $\Xi = \mathbb{R}$ и ε_t – процесс гауссовского $\mathcal{N}(0,1)$ белого шума. Авторегрессионный процесс $(X_t)_{t\in T}$ определяется следующим образом: $X_0 = x_0$, а далее X_t подчиняется закону*

$$X_{t+1} = \alpha X_t + \varepsilon_{t+1}.$$

Более общие модели авторегрессии задаются путём добавления дополнительных лагов $X_{t-1}, X_{t-2},..$ с другими коэффициентами (весами) $\alpha_{-1}, \alpha_{-2}, \dots$ В таком случае процесс $AR(p)$ задаётся рекуррентным уравнением

$$X_{t+1} = \sum_{i=0}^{p-1} \alpha_{-i} X_{t-i} + \varepsilon_{t+1}.$$

В теории временных рядов есть несколько методов, позволяющих оценивать веса лагов на основе реализации части траектории процесса X.

До сих пор мы рассматривали простые процессы: те, которые можно задать с помощью нескольких случайных величины и процесса белого шума, который в свою очередь, сам задаётся с помощью семейства одномерных распределений и требования независимости сечений. Один из известных способов задания более общего случайного процесса – задание семейства конечномерных распределений всех порядков. Инструмент для этого нам даёт известная теорема Колмогорова о существовании.

Теорема 1.1. *(Теорема Колмогорова о существовании) Пусть $(\Xi, \mathcal{G})$ – измеримое пространство, T – некоторое множество и для всех $k \in \mathbb{Z}_+$ и $t_1,..,t_k \in T$ заданы вероятностная мера $\nu_{t_1,..,t_k}$ на Ξ^k такая, что выполнены условия согласованности:*

1. *Для любого набора измеримых множеств $F_1,..,F_k \in \mathcal{G}$ и любой перестановки σ на множестве $\{1,..,k\}$ выполнено*

$$\nu_{t_1,..,t_k}(F_1 \times .. \times F_k) = \nu_{t_{\sigma(1)},..,t_{\sigma(k)}}\left(F_{\sigma(1)} \times .. \times F_{\sigma(k)}\right).$$

2. *Для всех $k, m \in \mathbb{Z}_+$ и $t_1, .., t_k, t_{k+1}, .., t_{k+m} \in T$ выполнено*

$$\nu_{t_1,..,t_k}(F_1 \times .. \times F_k) = \nu_{t_1,..,t_{k+m}}(F_1 \times .. \times F_k \times \Xi \times .. \times \Xi),$$

где справа число множителей равно $k + m$.

Тогда существует вероятностное пространство $(\Omega, \mathcal{F}, P)$ и случайный процесс $(X_t)_{t\in T}$, определённый на нём, такой, что

$$\nu_{t_1,..,t_k}(F_1 \times .. \times F_k) = P\left(X_{t_1} \in F_1, .., X_{t_k} \in F_k\right)$$

для всех $k \in \mathbb{Z}_+$, индексов $t_1, .., t_k \in T$ и измеримых множеств $F_1, .., F_k \in \mathcal{G}$.

Так, разумным образом строя семейства конечномерных распределений, мы можем корректно задать случайный процесс, как функцию $X : T \times \Omega \to \Xi$ на вероятностном пространстве $(\Omega, \mathcal{F}, P)$.

1.3 Пример: гауссовский процесс (отнесено в семинар)

Возможно, вы уже встречали гауссовские процессы в курсах по машинному обучению или анализу данных. Они являются очень мощным средством для разного рода моделирования и регрессии. Для Python имплементацию можно найти, например, в пакетах `scikit-learn` и `gpy`, ещё есть `GPyTorch`, как минимум одно решение на PyTorch.

Для простоты рассмотрим случай $T = \mathbb{R},\ \Xi = \mathbb{R}$; впрочем, если вас не пугают тензоры, вы можете подумать, как обобщить (например, вот так). Любой гауссовский процесс задаётся двумя функциями $m : T \to \mathbb{R}$ и $K : T \times T \to \mathbb{R}$, причём последняя должна быть симметричной и удовлетворять следующему условию.

Определение 1.3. *Функцию $K : T \times T \to \mathbb{R}$ назовём положительно полуопределённой, если для любого набора $t_1, .., t_k \in T$ и любых ненулевых векторов $x \in \mathbb{R}^k$*

$$x^\top B x \geq 0,$$

где матрица $B \in \mathbb{R}^{k\times k}$ имеет элементы $B_{ij} = K(t_i, t_j)$

Иными словами, мы требуем, чтобы любая матрица B, составленная как выше, была ковариационной матрицей случайного вектора для любого набора индексов.

Мы задаём *гауссовский процесс* $(X_t)_{t\in T}$ с матожиданием $m : T \to \mathbb{R}$ и ковариационной функцией $K : T\times T \to \mathbb{R}$ полагая, что для любого набора индексов $t_1, .., t_k \in T$ случайный вектор $[X_{t_1}, .., X_{t_k}]^\top$ имеет гауссовское распределение с матожиданием $[m(t_1), .., m(t_k)]$ и ковариационной матрицей $B = (K(t_i, t_j))_{j=1,..,k}$. Набор гауссовских распределений удовлетворяет условиям согласованности (проверьте!) в Теореме 1.1, следовательно, существует и случайный процесс с конечномерными гауссовскими распределениями, заданными выше.

В качестве простейшей ковариационной функции можно использовать, например, линейную или квадратично-экспоненциальную функцию с параметром $l \in \mathbb{R}$ (который можно настраивать методом максимального правдоподобия в задаче регрессии):

$$K(t,s) = ts, \quad K(t,s) = e^{-\frac{(t-s)^2}{2l^2}}. \tag{1.3}$$

Упражнение 1.2. *Как выглядят функции $K_{wn}(t,s), m(t)$, которые задают процесс гауссовского белого шума из Примера 1.2?*

Упражнение 1.3. *Приведите пример функции $K : T \times T \to \mathbb{R}$, которая является симметричной, но не положительно полуопределённой.*

Упражнение 1.4. *Приведите пример функции $K : T \times T \to \mathbb{R}$, которая является несимметричной, но положительно полуопределённой.*

Гауссовские процессы – очень гибкий инструмент, потому что можно использовать самые разные ковариационные функции и моделировать очень сложные зависимости.

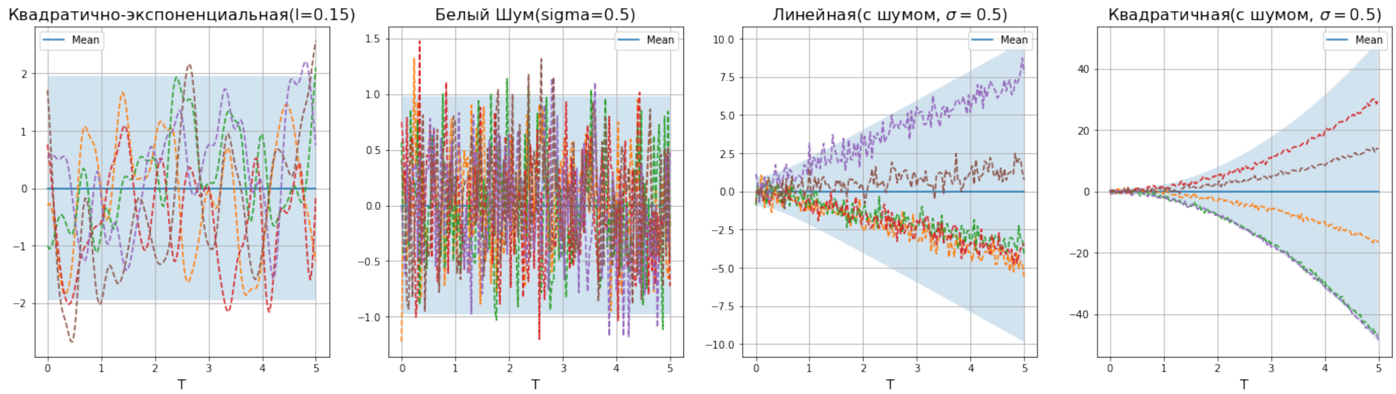

Рис. 1.4: Примеры траекторий гауссовских процессов с разными ковариациоными функциями: квадратично-экспоненциальной, белошумной $K_{wn}(t,s)$, линейной с шумом $K(t,s) = ts + K_{wn}(t,s)$ и квадратичной с шумом $K(t,s) = (ts)^2 + K_{wn}(t,s)$

1.4 Числовые характеристики и стационарность

Пусть процесс X принимает значения в $\Xi \subset \mathbb{R}$ или $\Xi \subset \mathbb{R}^n$. Мы уже познакомились с двумя характеристиками, когда обсуждали гауссовские процессы. Определим их явно и добавим ещё несколько.

Определение 1.4. *Пусть $(X_t)_{t \in T}$ – случайный процесс со значениями в $\mathbb{R}$ или $\mathbb{R}^n$.*

Функция $m(t) := \mathbb{E}[X_t]$ *называется матожиданием процесса* X.

Функция $K(t,s) := \mathrm{Cov}(X_t, X_s) = \mathbb{E}\left[(X_t - m(t))(X_s - m(s))^\top\right]$ *называется ковариационной функцией.*

Функция $v(t) := \mathrm{Tr}(K(t,t))$, *где* Tr *– это след матрицы, называется дисперсией процесса* X.

Упражнение 1.5. *Чему будут равны матожидания и ковариационные функции процессов из примеров 1.1, 1.3?*

В случае гауссовских процессов мы задали целый случайный процесс используя только матожидание и ковариационную функцию. Причём оказывается, что свойство положительной полуопределённости (Опр. 1.3) является существенным.

Утверждение 1.2. *(Упражнение) Функция* $K(t,s)$ *является ковариационной функцией некоторого процесса тогда и только тогда, когда она является симметричной и положительно полуопределённой в смысле Опр. 1.3.*

Пусть $(X_t)_{t\in T}$ – случайный процесс, определённый на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и с векторным пространством T в качестве индексного (чтобы мы могли делать сдвиги на элемент $s \in T$).

Определение 1.5. *Процесс* X *называется стационарным в узком смысле (strongly stationary), если для любого конечного набора индексов* $t_1, .., t_k$ *конечномерные распределения не меняются со сдвигом всего набора на* $s \in T$, *то есть,*

$$P\left(X_{t_1} \in A_1, .., \ X_{t_k} \in A_k\right) = P\left(X_{t_1+s} \in A_1, .., \ X_{t_k+s} \in A_k\right)$$

для любых измеримых множеств $A_1, .., A_k \in \mathcal{G}$.

Определение 1.6. *Процесс* X *называется стационарным в широком смысле (weakly stationary), если*

1. *Существуют конечные матожидание* $m(t)$ *и дисперсия* $v(t)$ *в каждой точке* $t \in T$;
2. *Матожидание – константа:* $m(t) = const$;
3. *Ковариационная фукнция зависит только от разности аргументов: существует функция* $\widetilde{K} : T \to \mathbb{R}$ *такая, что* $K(t,s) = \widetilde{K}(|t-s|) < \infty$ *для любых* $t, s \in T$.

Пример 1.4. *В общем случае из стационарности в узком смысле следует стационарность в широком смысле (докажите, для этого нужно выписать ковариационные функции* $K(t, t+s)$ *и* $K(t-s, t)$ *и показать, что они равны для любых* $t, s \in T$, *используя*

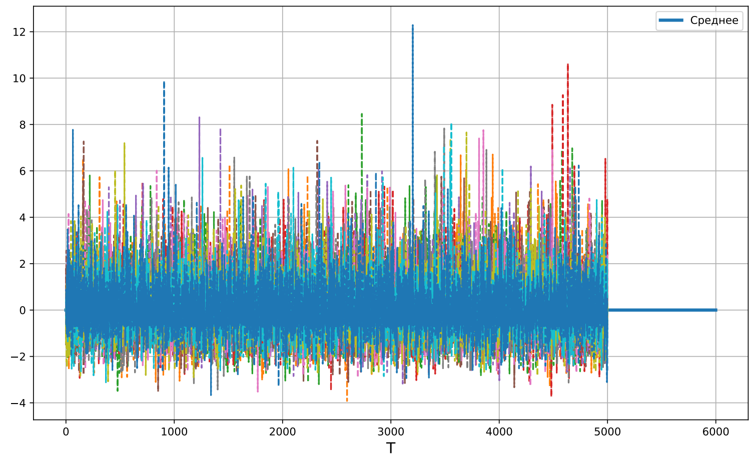

Рис. 1.5: Траектории процесса из Примера 1.4

определение стационарности в узком смысле). В обратную сторону, как оказывается, утверждение неверно – существуют процессы, которые стационарны в широком смысле, но не стационарны в узком. Пусть $T = \mathbb{Z}$, $\Xi = \mathbb{R}$, ε_t – процесс белого шума с распределением $\mathcal{N}(0, 1)$. Процесс X_t, определённый как

$$X_t = \begin{cases} \varepsilon_t, & t \text{ – чётное}, \\ \frac{1}{\sqrt{2}}(\varepsilon_t^2 - 1), & t \text{ – нечётное}, \end{cases}$$

не будет стационарным в узком смысле, но будет стационарным в широком смысле.

Пример 1.5. *Гауссовские процессы интересны ещё и тем, что для них стационарность в широком смысле влечёт стационарность в узком смысле (проверьте, это следует напрямую из того, что совместные распределения всех конечных наборов сечений гауссовские).*

Упражнение 1.6. *Проверьте, являются ли процесса из Примера 1.1 и гауссовский белый шум из Примера 1.2 стационарными в узком или широком смыслах?*

Упражнение 1.7. *Пусть $T \in \mathbb{R}$, $(X_t)_{t \in T}$ – гауссовский процесс с $m(t) = 0$. Проверьте, будет ли процесс стационарным (в широком смысле), если ковариацонная функция*

1. *линейная $K(t, s) = ts$;*
2. *квадратично-экспоненциальная $K(t, s) = e^{-\frac{(t-s)^2}{2l^2}}$ с параметром $l \in \mathbb{R}$.*

Литература

[1] Alfred Cowles 3rd. Can stock market forecasters forecast? *Econometrica*, 1(3):309–324, 1933.

[2] Alfred Cowles 3rd and Herbert E. Jones. Some a posteriori probabilities in stock market action. *Econometrica*, 5(3):280–294, 1937.

[3] Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 3e série, 17:21–86, 1900.

[4] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[5] Michael Brin and Garrett Stuck. *Introduction to Dynamical Systems*. Cambridge University Press, 2002.

[6] Yuansi Chen, Raaz Dwivedi, Martin J. Wainwright, and Bin Yu. Fast mixing of metropolized hamiltonian monte carlo: benefits of multi-step gradients. *J. Mach. Learn. Res.*, 21(1), jan 2020.

[7] Persi Diaconis. The markov chain monte carlo revolution. *Bulletin of the American Mathematical Society*, 46:179textendash205, 04 2009.

[8] Alain Durmus and Éric Moulines. High-dimensional bayesian inference via the unadjusted langevin algorithm. *Bernoulli*, 2016.

[9] Raaz Dwivedi, Yuansi Chen, Martin J Wainwright, and Bin Yu. Log-concave sampling: Metropolis-hastings algorithms are fast! In Sébastien Bubeck, Vianney Perchet, and Philippe Rigollet, editors, *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 793–797. PMLR, 06–09 Jul 2018.

[10] Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[11] L. Isserlis. On a Formula for the Product-Moment Coefficient of any Order of a Normal Frequency Distribution in any Number of Variables, November 1918.

[12] Galin Jones. On the markov chain central limit theorem. *Probability Surveys*, 1, 09 2004.

[13] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960.

[14] M. G. Kendall and A. Bradford Hill. The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1):11–34, 1953.

[15] Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144, 1976.

[16] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 06 1953.

[17] Bernt Oksendal. *Stochastic Differential Equations (3rd Ed.): An Introduction with Applications.* Springer-Verlag, Berlin, Heidelberg, 1992.

[18] Paul A. Samuelson. Proof that properly discounted present values of assets vibrate randomly. *The Bell Journal of Economics and Management Science*, 4(2):369–374, 1973.

[19] O. A. Stepanov. Kalman filtering: Past and present. an outlook from russia. (on the occasion of the 80th birthday of rudolf emil kalman). *Gyroscopy and Navigation*, 2(2):99–110, Apr 2011.

[20] R. L. Stratonovich. Conditional markov processes. *Theory of Probability & Its Applications*, 5(2):156–178, 1960.

[21] Holbrook Working. A random-difference series for use in the analysis of time series. *Journal of the American Statistical Association*, 29(185):11–24, 1934.

[22] L.E. Zachrisson. On optimal smoothing of continuous time kalman processes. *Information Sciences*, 1(2):143–172, 1969.

[23] Guangyao Zhou. Metropolis augmented hamiltonian monte carlo. In *Fourth Symposium on Advances in Approximate Bayesian Inference*, 2022.

[24] Н. В. Рекнер М. Шапошников С. В. Богачев, В. И. Крылов. *Уравнения Фоккера – Планка – Колмогорова.* Институт компьютерных исследований, 2013.

[25] Ширяев А.Н. Булинский А.В. *Теория случайных процессов.* М:ФИЗМАТЛИТ, 2005.

[26] Тиморин В.А. *Геометрия гамильтоновых систем и уравнений с частными производными.* Москва : ВШЭ, 2017.

[27] Синай Я.Г. Коралов Л.Б. *Теория вероятностей и случайные процессы.* МЦНМО, 2013.

[28] Р.Л. Стратонович. *Условные марковские процессы и их применение к теории оптимального управления.* Московский государственный университет, 1966.

[29] А.Н. Ширяев. *Основы стохастической финансовой математики.* МЦНМО, 2016.