Лекция 2
Винеровский и пуассоновский процессы. Генерирование реализаций случайных процессов

Артемов А. В.
бПМИ ФКН ВШЭ

27 января 2018

1 Винеровский и пуассоновский процессы

1.1 Лирическое отступление: гауссовские векторы

Сделаю небольшое лирическое отступление и вспомним гауссовские векторы. Для этого вспомним, что такое *характеристическая функция* случайной величины и вектора.

Определение 1. Пусть ξ — случайная величина с плотностью p_ξ. Тогда её характеристической функцией называется функция $\varphi_\xi : \mathbb{R} \mapsto \mathbb{C}$, определяемая следующим образом:

$$\varphi_\xi(t) = \mathsf{E}\big[e^{it\xi}\big] = \int\limits_{-\infty}^{+\infty} e^{itx} p_\xi(x)\, dx.$$

Определение 2. Пусть $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ — случайный вектор с совместной плотностью $p_{\boldsymbol{\xi}}$. Тогда её характеристической функцией называется функция $\varphi_{\boldsymbol{\xi}} : \mathbb{R}^n \mapsto \mathbb{C}$, определяемая следующим образом: $\varphi_{\boldsymbol{\xi}}(\mathbf{t}) = \mathsf{E}\big[e^{i\langle \boldsymbol{\xi}, \mathbf{t}\rangle}\big]$, где $\langle \cdot, \cdot \rangle$ — скалярное произведение в $\mathbb{R}^n$.

По сути, характеристическая функция — это преобразование Фурье функции распределения.

Далее, из курса теории вероятности известно, что если $\xi \sim \mathcal{N}(\mu, \sigma^2)$, то

$$\varphi_\xi(t) = \exp\left\{ i\mu t - \frac{1}{2}\sigma^2 t^2 \right\}.$$

Поэтому гауссовский вектор вводят следующим образом:

Определение 3. Случайный вектор $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ подчиняется *многомерному нормальному распределению*, если его характеристическая функция равна

$$\varphi_{\boldsymbol{\xi}}(\mathbf{t}) = \exp\left\{ i\langle \boldsymbol{\mu}, \mathbf{t}\rangle - \frac{1}{2}\langle \boldsymbol{\Sigma}\mathbf{t}, \mathbf{t}\rangle \right\},$$

где $\boldsymbol{\mu} \in \mathbb{R}^n$ — некоторый фиксированный вектор, а $\boldsymbol{\Sigma}$ — некоторая симметрическая и неотрицательно определённая матрица. В таком случае пишут, что $\xi \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Это определение не очень удобно. Докажем одну теорему, которая даст несколько более удобное определение.

Теорема 1. *Пусть $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ — случайный вектор. Он будет гауссовским тогда и только тогда, когда для любого неслучайного вектора $\boldsymbol{\lambda} \in \mathbb{R}^n$ $\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle$ имеет нормальное распределение.*

Доказательство. Пусть $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Тогда посмотрим на характеристическую функцию $\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle$. Заметим, что она равна

$$\varphi_{\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle}(t) = \mathsf{E}\big[e^{it\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle}\big] = \mathsf{E}\big[e^{i\langle t\boldsymbol{\lambda}, \boldsymbol{\xi} \rangle}\big] = \varphi_{\boldsymbol{\xi}}(t\boldsymbol{\lambda}) = \exp\left\{ it\langle \boldsymbol{\mu}, \boldsymbol{\lambda} \rangle - \frac{t^2}{2}\langle \boldsymbol{\Sigma}\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle \right\}.$$

Это означает, что $\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle \sim \mathcal{N}(\langle \boldsymbol{\mu}, \boldsymbol{\lambda} \rangle, \langle \boldsymbol{\Sigma}\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle)$.

Теперь предположим, что $\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle$ имеет нормальное распределение для любого $\boldsymbol{\lambda}$. Тогда посмотрим на характеристическую функцию $\boldsymbol{\xi}$:

$$\varphi_{\boldsymbol{\xi}}(\boldsymbol{\lambda}) = \mathsf{E}\big[e^{i\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle}\big] = \varphi_{\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle}(1) = \exp\left\{ i\,\mathsf{E}[\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle] - \frac{1}{2}\mathsf{D}[\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle] \right\}.$$

Теперь заметим, что

$$\mathsf{E}[\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle] = \mathsf{E}\left[\sum_{k=1}^{n} \lambda_k \xi_k\right] = \sum_{k=1}^{n} \lambda_k\,\mathsf{E}[\xi_k] = \langle \boldsymbol{\lambda}, \mathsf{E}[\boldsymbol{\xi}] \rangle, \tag{1.1}$$

$$\mathsf{D}[\langle \boldsymbol{\lambda}, \boldsymbol{\xi} \rangle] = \mathsf{cov}\left(\sum_{k=1}^{n} \lambda_k \xi_k, \sum_{k=1}^{n} \lambda_k \xi_k\right) = \sum_{i,j=1}^{n} \mathsf{cov}(\xi_i, \xi_j)\lambda_i\lambda_j = \langle \mathsf{D}[\boldsymbol{\xi}]\boldsymbol{\lambda}, \boldsymbol{\lambda} \rangle. \tag{1.2}$$

Отсюда получаем, что $\boldsymbol{\xi} \sim \mathcal{N}(\mathsf{E}[\boldsymbol{\xi}], \mathsf{D}[\boldsymbol{\xi}])$. □

Теперь выпишем следствия из этой теоремы.

Следствие (Смысл параметров)**.** *Если случайный вектор $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, то $\boldsymbol{\mu} = \mathsf{E}[\xi]$, $\boldsymbol{\Sigma} = \mathsf{D}[\xi]$.*

Следствие (Линейные преобразования)**.** *Любое линейное преобразование гауссовского вектора тоже является гауссовским вектором.*

Теперь вспомним про плотности. У гауссовских векторов она есть не всегда.

Теорема 2 (о плотности гауссовских векторов)**.** *Пусть $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ — n-мерный гауссовский вектор. Тогда, если $\boldsymbol{\Sigma}$ положительно определена, то существует плотность $p_{\boldsymbol{\xi}}(\mathbf{t})$ и она равна*

$$p_{\boldsymbol{\xi}}(\mathbf{t}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left\{ -\frac{1}{2}\left\langle \boldsymbol{\Sigma}^{-1}(\mathbf{t} - \boldsymbol{\mu}), (\mathbf{t} - \boldsymbol{\mu}) \right\rangle \right\}.$$

Задача 1. *Пусть $\mathbf{X} = (\xi, \eta)$ — гауссовский вектор, для которого:*[1]

$$\mathsf{E}[\mathbf{X}] = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \mathsf{D}[\mathbf{X}] = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad |\rho| < 1.$$

Докажите, что плотность случайного вектора $\mathbf{X}$ равна

$$p_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Bigg\{ -\frac{1}{2(1-\rho^2)}\Bigg(\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 - 2\rho\frac{x_1 - \mu_1}{\sigma_1}\frac{x_2 - \mu_2}{\sigma_2}\Bigg)\Bigg\}$$

[1]Несложно понять, что ρ есть коэффициент корреляции между ξ и η.

Порой хочется сказать, что любой вектор, состоящий из нормальных случайных величин, является гауссовским. Но это неверно.

Пример 1. Пусть ξ_1 и ξ_2 — это независимые стандартные нормальные случайные величины. Построим случайный вектор (X_1, X_2) следующим образом:

$$(X_1, X_2) = \begin{cases} (\xi_1, |\xi_2|), & \xi_1 \geqslant 0 \\ (\xi_1, -|\xi_2|), & \xi_1 < 0 \end{cases}$$

Данный случайный вектор не будет гауссовским (почему?).

У гауссовских векторов есть одно уникальное свойство, связанное с некоррелированностью.

Теорема 3. *Пусть $\boldsymbol{\xi}$ — гауссовский вектор. Тогда его компоненты независимы тогда и только тогда, когда они некоррелированны.*

Вопрос: допустим, что у нас есть вектор, состоящий из некоррелированных нормальных случайных величин. Можно ли сказать, что он гауссовский? Оказывается, что нет.

Пример 2. Пусть $X \sim \mathcal{N}(0, 1)$ и $c \geqslant 0$. Построим по ней новую случайную величину Y следующим образом:

$$Y = \begin{cases} X, & |X| \leqslant c \\ -X, & |X| > c \end{cases}$$

Оказывается, что $Y \sim \mathcal{N}(0, 1)$. Тогда $\mathsf{cov}(X, Y) = \mathsf{E}[XY]$, что, в свою очередь, равно $\mathsf{E}[X^2\, \mathrm{I}\{|X| \leqslant c\}] - \mathsf{E}[X^2\, \mathrm{I}\{|X| > c\}]$. Что мы можем сказать про ковариацию?

- Она является непрерывной функцией от c.
- Если $c = 0$, то $\mathsf{cov}(X, Y) = -1$, а если $c = +\infty$, то $\mathsf{cov}(X, Y) = 1$.

В таком случае можно сказать, что есть c такая, что $\mathsf{cov}(X, Y) = 0$. Зафиксируем её. Можно ли сказать, что (X, Y) — это гауссовский вектор? Увы, но нет. Если бы это было так, то X и Y были бы независимы. Но $\mathsf{P}(X > c, Y > c) = 0 \neq \mathsf{P}(X > c)\, \mathsf{P}(Y > c)$.

Теперь расскажем два факта, которые могут понадобиться в дальнейшем и связаны с нормальным распределением.

Пример 3. Пусть (ξ, η) — случайный вектор с совместной плотностью

$$p_{\xi,\eta}(x, y) = C \exp\{-(1 + x^2)(1 + y^2)\}, \text{ где } C = \iint\limits_{\mathbb{R}^2} \exp\{-(1 + x^2)(1 + y^2)\}\, dx\, dy$$

Попробуем найти условную плотность $p_{\xi|\eta}(x \mid y)$. Сразу же заметим, что процесс аналогичен для $p_{\eta|\xi}(y \mid x)$. Для этого найдём плотность η:

$$p_\eta(y) = \int\limits_{-\infty}^{+\infty} Ce^{-(1+x^2)(1+y^2)}\, dx = Ce^{-(1+y^2)} \int\limits_{-\infty}^{+\infty} e^{-x^2(1+y^2)}\, dx =$$

$$= \frac{Ce^{-(1+y^2)}}{\sqrt{y^2+1}} \int\limits_{-\infty}^{+\infty} e^{-u^2}\, du = \frac{C\sqrt{\pi}e^{-(1+y^2)}}{\sqrt{y^2+1}}.$$

Теперь несложно посчитать условную плотность:

$$p_{\xi|\eta}(x \mid y) = \frac{g(x, y)}{p_\eta(y)} = \frac{\sqrt{y^2+1}}{\sqrt{\pi}} e^{-x^2(1+y^2)} = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{x^2}{2\sigma_y^2}}, \text{ где } \sigma_y^2 = \frac{1}{2 + 2y^2}.$$

Из математической статистики вам известны такие понятия, как оценки. Следующая теорема в некоторой степени связана с ними, но это станет ясней позднее, когда дело дойдёт до фильтров Калмана.

Теорема 4 (о нормальной корреляции, одномерный случай). *Пусть (ξ, η) — двумерный гауссовский вектор. Тогда*

$$\mathsf{E}[\eta \mid \xi] = \mathsf{E}[\eta] + \frac{\mathsf{cov}(\xi, \eta)}{\mathsf{D}[\xi]}(\xi - \mathsf{E}[\xi]), \tag{1.3}$$

$$\Delta = \mathsf{E}\big[(\eta - \mathsf{E}[\eta \mid \xi])^2\big] = \mathsf{D}[\eta] - \frac{\mathsf{cov}^2(\xi, \eta)}{\mathsf{D}[\xi]}. \tag{1.4}$$

Доказательство. Докажем эту теорему в лоб. Для этого посчитаем условную плотность $p_{\eta|\xi}(y \mid x)$. Формула для совместной плотности была выведена в задаче 1. Пользуясь теми же обозначениями, получим (проверьте!), что

$$p_{\eta|\xi}(y \mid x) = \frac{1}{\sqrt{2\pi\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left(y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2\right\}.$$

Из этого можно сделать следующий вывод:

$$(\eta \mid \xi = x) \sim \mathcal{N}\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1-\rho^2)\right) \implies \mathsf{E}[\eta \mid \xi] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(\xi - \mu_1).$$

Заменяя, получим первую часть утверждения. Для доказательства второго утверждения подставим полученный результат и раскроем скобки как квадрат разности:

$$\Delta = \mathsf{E}\left[\left(\eta - \mathsf{E}[\eta] - \frac{\mathsf{cov}(\xi,\eta)}{\mathsf{D}[\xi]}(\xi - \mathsf{E}[\xi])\right)^2\right] = \mathsf{E}\big[(\eta - \mathsf{E}[\eta])^2\big] -$$
$$- 2\frac{\mathsf{cov}(\xi,\eta)}{\mathsf{D}[\xi]}\mathsf{E}[(\eta - \mathsf{E}[\eta])(\xi - \mathsf{E}[\xi])] + \frac{\mathsf{cov}^2(\xi,\eta)}{\mathsf{D}^2[\xi]}\mathsf{E}\big[(\xi - \mathsf{E}[\xi])^2\big] = \mathsf{D}[\eta] - \frac{\mathsf{cov}^2(\xi,\eta)}{\mathsf{D}[\xi]}. \quad \square$$

У данной теоремы есть многомерное обобщение.

Теорема 5 (о нормальной корреляции, общий случай). *Пусть $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ — n-мерный гауссовский вектор. Сделаем следующее разбиение:*

$$\mathbf{X} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} \text{ с размерами } \begin{pmatrix} q \times 1 \\ (n-q) \times 1 \end{pmatrix}$$

Соответственным образом вводятся разбиения матожидания и матрицы ковариаций:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \text{ с размерами } \begin{pmatrix} q \times 1 \\ (n-q) \times 1 \end{pmatrix} \tag{1.5}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \text{ с размерами } \begin{pmatrix} q \times q & q \times (n-q) \\ (n-q) \times q & (n-q) \times (n-q) \end{pmatrix} \tag{1.6}$$

Тогда верны следующие формулы:

$$\mathsf{E}[\boldsymbol{\xi} \mid \boldsymbol{\eta}] = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{\eta} - \boldsymbol{\mu}_2), \tag{1.7}$$

$$\Delta = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}. \tag{1.8}$$

Доказательство. Для начала найдём линейную комбинацию $\boldsymbol{\delta} = \boldsymbol{\xi} + \mathbf{A}\boldsymbol{\eta}$ такую, что она некоррелирована (а следовательно, и независима) с $\boldsymbol{\eta}$. Для этого распишем корреляцию:

$$\mathsf{cov}(\boldsymbol{\delta}, \boldsymbol{\eta}) = \mathsf{cov}(\boldsymbol{\xi}, \boldsymbol{\eta}) + \mathsf{cov}(\mathbf{A}\boldsymbol{\eta}, \boldsymbol{\eta}) = \mathsf{cov}(\boldsymbol{\xi}, \boldsymbol{\eta}) + \mathbf{A}\,\mathsf{D}[\boldsymbol{\eta}] = \boldsymbol{\Sigma}_{12} + \mathbf{A}\boldsymbol{\Sigma}_{22} \implies \mathbf{A} = -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}.$$

Отсюда получаем, что

$$\begin{aligned}\mathsf{E}[\boldsymbol{\xi} \mid \boldsymbol{\eta}] = \mathsf{E}[\boldsymbol{\delta} - \mathbf{A}\boldsymbol{\eta} \mid \boldsymbol{\eta}] = \mathsf{E}[\boldsymbol{\delta} \mid \boldsymbol{\eta}] - \mathsf{E}[\mathbf{A}\boldsymbol{\eta} \mid \boldsymbol{\eta}] = \mathsf{E}[\boldsymbol{\delta}] - \mathbf{A}\boldsymbol{\eta} = \boldsymbol{\mu}_1 + \mathbf{A}(\boldsymbol{\mu}_2 - \boldsymbol{\eta}) = \\ = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{\eta} - \boldsymbol{\mu}_2).\end{aligned}$$

Теперь посмотрим на то, как себя ведёт Δ. Как известно, по формуле полной вероятности:

$$\Delta = \mathsf{E}\big[(\boldsymbol{\xi} - \mathsf{E}[\boldsymbol{\xi} \mid \boldsymbol{\eta}])^2\big] = \mathsf{E}\big[\mathsf{E}\big[(\boldsymbol{\xi} - \mathsf{E}[\boldsymbol{\xi} \mid \boldsymbol{\eta}])^2 \bigm| \boldsymbol{\eta}\big]\big] = \mathsf{E}[\mathsf{D}[\boldsymbol{\xi} \mid \boldsymbol{\eta}]].$$

Теперь посмотрим на условную дисперсию:

$$\mathsf{D}[\boldsymbol{\xi} \mid \boldsymbol{\eta}] = \mathsf{D}[\boldsymbol{\delta} - \mathbf{A}\boldsymbol{\eta} \mid \boldsymbol{\eta}] = \mathsf{D}[\boldsymbol{\delta} \mid \boldsymbol{\eta}] + \mathsf{D}[\mathbf{A}\boldsymbol{\eta} \mid \boldsymbol{\eta}] - \mathsf{cov}(\boldsymbol{\delta}, \mathbf{A}\boldsymbol{\eta} \mid \boldsymbol{\eta}) - \mathsf{cov}(\mathbf{A}\boldsymbol{\eta}, \boldsymbol{\delta} \mid \boldsymbol{\eta}).$$

Как известно, матожидание вектора, умноженного на матрицу слева/справа, равно матожиданию вектора, умноженного на эту же матрицу слева/справа, а условная дисперсия случайной величины, не зависящей от условия, равна обычной дисперсии. Тогда это равно

$$\mathsf{D}[\boldsymbol{\delta} \mid \boldsymbol{\eta}] + \mathbf{A}\,\mathsf{D}[\boldsymbol{\eta} \mid \boldsymbol{\eta}]\mathbf{A}^\intercal - \mathsf{cov}(\boldsymbol{\delta}, \boldsymbol{\eta} \mid \boldsymbol{\eta})\mathbf{A}^\intercal - \mathbf{A}\,\mathsf{cov}(\boldsymbol{\eta}, \boldsymbol{\delta} \mid \boldsymbol{\eta}) = \mathsf{D}[\boldsymbol{\delta} \mid \boldsymbol{\eta}] = \mathsf{D}[\boldsymbol{\delta}].$$

Осталось посчитать эту дисперсию. Для этого распишем дисперсию суммы, пользуясь свойствами матриц $\boldsymbol{\Sigma}_{ij}$:

$$\begin{aligned}\mathsf{D}[\boldsymbol{\delta}] &= \mathsf{D}[\boldsymbol{\xi} + \mathbf{A}\boldsymbol{\eta}] = \mathsf{D}[\boldsymbol{\xi}] + \mathsf{D}[\mathbf{A}\boldsymbol{\eta}] + \mathsf{cov}(\boldsymbol{\xi}, \mathbf{A}\boldsymbol{\eta}) + \mathsf{cov}(\mathbf{A}\boldsymbol{\eta}, \boldsymbol{\xi}) = \\ &= \mathsf{D}[\boldsymbol{\xi}] + \mathbf{A}\,\mathsf{D}[\boldsymbol{\eta}]\mathbf{A}^\intercal + \mathsf{cov}(\boldsymbol{\xi}, \boldsymbol{\eta})\mathbf{A}^\intercal + \mathbf{A}\,\mathsf{cov}(\boldsymbol{\eta}, \boldsymbol{\xi}) = \\ &= \boldsymbol{\Sigma}_{11} + (-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{22}(-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})^\intercal + \boldsymbol{\Sigma}_{12}(-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})^\intercal + (-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{21} = \\ &= \boldsymbol{\Sigma}_{11} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = \\ &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\end{aligned}$$

Так как матожидание константы есть сама константа, то $\Delta = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. □

1.2 Примеры случайных процессов

Наше небольшое введение закончилось. Теперь можно посмотреть на несколько основных примеров случайных процессов.

1.2.1 Гауссовский и винеровский процессы

Многие процессы, которые попадаются на практике, обладают так называемыми независимыми приращениями. Что это значит?

Определение 4. Пусть $X = (X_t)_{t \geqslant 0}$ — некоторый случайный процесс. Будем говорить, что X есть *процесс с независимыми приращениями*, если для любых $t_0, t_1, \ldots, t_n \in T$ таких, что $0 = t_0 < t_1 < \ldots < t_n$ случайные величины X_{t_0}, $X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ независимы в совокупности.

В 1827 году Роберт Броун открыл движение пыльцевых зёрен в жидкости. Исследуя пыльцу под микроскопом, он установил, что в растительном соке плавающие пыльцевые зёрна двигаются совершенно хаотически зигзагообразно во все стороны. В дальнейшем это хаотическое движение назвали *броуновским*. Для его математического описания используется так называемый *винеровский процесс*. Как он вводится?

Определение 5. Случайный процесс $B = (B_t)_{t \geqslant 0}$ называется винеровским, если для него выполнены следующие условия:

1. $B_0 = 0$ почти наверное.
2. B — процесс с независимыми приращенимы.
3. $B_t - B_s \sim \mathcal{N}(0, t - s)$ для любых $0 \leqslant s < t < +\infty$.
4. B имеет непрерывные почти наверное траектории, то есть с вероятностью 1 B_t непрерывна, как функция от t.

Обычно наряду с винеровскими процессами вводят *гауссовские процессы*.

Определение 6. Случайный процесс $X = (X_t)_{t \in T}$ называется гауссовским, если все его конечномерные функции распределения являются гауссовскими, то есть задают гауссовский вектор.

Свойство 1. *Гауссовский процесс однозначно определяется своим математическим ожиданием и ковариационной функцией.*

Пример 4. Оказывается, что винеровский процесс является гауссовским. Действительно, возьмём произвольные $t_0, t_1, \ldots, t_n$ таким образом, что $0 = t_0 < t_1 < \cdots < t_n$. Как известно, у винеровского процесса независимые приращения. Следовательно, $B_{t_0}, B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ независимы в совокупности и образуют гауссовский вектор. Теперь поймём, какие распределения имеют $B_{t_0}, B_{t_1}, \ldots, B_{t_n}$. Для этого заметим, что

$$\begin{pmatrix} B_{t_0} \\ B_{t_1} \\ B_{t_2} \\ \vdots \\ B_{t_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} B_{t_0} \\ B_{t_1} - B_{t_0} \\ B_{t_2} - B_{t_1} \\ \vdots \\ B_{t_n} - B_{t_{n-1}} \end{pmatrix}$$

Следовательно, $B_{t_i} \sim \mathcal{N}(0, t_i)$ и они образуют гауссовский вектор.

Из этого сразу же получаем, что $\mathsf{E}[B_t] = 0$. Теперь покажем, что $R_B(t_1, t_2) = \min(t_1, t_2)$. Без ограничения общности будем считать, что $t_1 < t_2$. Тогда

$$R_B(t_1, t_2) = \mathsf{E}[B_{t_1} B_{t_2}] = \mathsf{E}[B_{t_1}((B_{t_2} - B_{t_1}) + B_{t_1})] = \mathsf{E}[B_{t_1}(B_{t_2} - B_{t_1})] + \mathsf{E}\big[B_{t_1}^2\big].$$

Так как B_{t_1} и $B_{t_2} - B_{t_1}$ независимы, то $\mathsf{E}[B_{t_1}(B_{t_2} - B_{t_1})] = \mathsf{E}[B_{t_1}]\,\mathsf{E}[B_{t_2} - B_{t_1}] = 0$. Тем самым мы получаем, что $R_B(t_1, t_2) = t_1$.

1.2.2 Процесс Орнштейна-Уленбека

Этот процесс пошёл из теории стохастических дифференциальных уравнений. Пусть $B = (B_t)_{t \geqslant 0}$ — винеровский процесс. Построим по нему новый процесс $X = (X_t)_{t \geqslant 0}$ по следующему правилу: $X_t = e^{-t} B_{e^{2t}}$. Полученный процесс называется *процессом Орнштейна-Уленбека.* Каковы его свойства?

Свойство 2. $\mathsf{E}[X_t] = 0$, $R_X(t,s) = e^{-|t-s|}$.

Доказательство. Первая часть очевидна: $\mathsf{E}[X_t] = \mathsf{E}[e^{-t} B_{e^{2t}}] = 0$. Теперь рассмотрим ковариационную функцию:

$$R_X(t,s) = \mathsf{E}[X_t X_s] = e^{-(s+t)}\,\mathsf{E}[B_{e^{2t}} B_{e^{2s}}] = e^{2\min(t,s)-(s+t)} = e^{-|t-s|}. \qquad \square$$

Свойство 3. *Процесс Орнштейна-Уленбека является гауссовским.*

Доказательство. Без ограничения общности зафиксируем числа $t_1 < t_2 < \ldots < t_n$. Рассмотрим случайный вектор $(X_{t_1}, \ldots, X_{t_n})$. Как он устроен? Распишем последний член:

$$\begin{aligned} X_{t_n} = e^{-t_n} B_{e^{2t_n}} = e^{-t_n}(B_{e^{2t_n}} - B_{e^{2t_{n-1}}}) + e^{-t_n} B_{e^{2t_{n-1}}} = \ldots = \\ = e^{-t_n} \sum_{k=1}^{n-1} (B_{e^{2t_{k+1}}} - B_{e^{2t_k}}) + e^{-t_n} B_{e^{2t_1}} \end{aligned}$$

Далее, нам известно, что $B_{e^{2t_1}}, B_{e^{2t_2}} - B_{e^{2t_1}}, \ldots, B_{e^{2t_n}} - B_{e^{2t_{n-1}}}$ независимы в совокупности и имеют следующие распределения:

$$B_{e^{2t_1}} \sim \mathcal{N}(0, e^{2t_1}), \quad B_{e^{2t_k}} - B_{e^{2t_{k-1}}} \sim \mathcal{N}(0, e^{2t_k} - e^{2t_{k-1}}).$$

Осталось заметить, что

$$\begin{pmatrix} X_{t_1} \\ X_{t_2} \\ X_{t_3} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} e^{-t_1} & 0 & 0 & \ldots & 0 \\ e^{-t_2} & e^{-t_2} & 0 & \ldots & 0 \\ e^{-t_3} & e^{-t_3} & e^{-t_3} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-t_n} & e^{-t_n} & e^{-t_n} & \ldots & e^{-t_n} \end{pmatrix} \begin{pmatrix} B_{e^{2t_1}} \\ B_{e^{2t_2}} - B_{e^{2t_1}} \\ B_{e^{2t_3}} - B_{e^{2t_2}} \\ \vdots \\ B_{e^{2t_n}} - B_{e^{2t_{n-1}}} \end{pmatrix}$$

Отсюда получаем, что $(X_{t_1}, \ldots, X_{t_n})$ — действительно гауссовский вектор. $\square$

Теперь посчитаем совместную плотность такого вектора. Для этого выпишем совместную плотность $\boldsymbol{\xi} = (B_{e^{2t_1}}, B_{e^{2t_2}} - B_{e^{2t_1}}, \ldots, B_{e^{2t_n}} - B_{e^{2t_{n-1}}})$. Так как компоненты независимы, то она равна произведению плотностей каждой компоненты:

$$p_{\boldsymbol{\xi}}(x_1, \ldots, x_n) = \frac{1}{\sqrt{2\pi} e^{t_1}} \exp\left\{-\frac{x_1^2}{2e^{2t_1}}\right\} \prod_{k=2}^{n} \frac{1}{\sqrt{2\pi(e^{2t_k} - e^{2t_{k-1}})}} \exp\left\{-\frac{x_k^2}{2(e^{2t_k} - e^{2t_{k-1}})}\right\}$$

Теперь выразим компоненты вектора $\boldsymbol{\xi}$ через компоненты вектора $\boldsymbol{\eta} = (X_1, X_2, \ldots, X_n)$:

$$\begin{cases} B_{e^{2t_1}} = e^{t_1} X_1 \\ B_{e^{2t_2}} = e^{t_2} X_2 \\ \ldots \\ B_{e^{2t_n}} = e^{t_n} X_n \end{cases} \Longrightarrow \begin{cases} B_{e^{2t_1}} = e^{t_1} X_1 \\ B_{e^{2t_2}} - B_{e^{2t_1}} = e^{t_2} X_2 - e^{t_1} X_1 \\ \ldots \\ B_{e^{2t_n}} - B_{e^{2t_{n-1}}} = e^{t_n} X_n - e^{t_{n-1}} X_{n-1} \end{cases}$$

В матричном виде это можно записать следующим образом:

$$\begin{pmatrix} B_{e^{2t_1}} \\ B_{e^{2t_2}} - B_{e^{2t_1}} \\ B_{e^{2t_3}} - B_{e^{2t_2}} \\ \vdots \\ B_{e^{2t_n}} - B_{e^{2t_{n-1}}} \end{pmatrix} = \begin{pmatrix} e^{t_1} & 0 & 0 & \dots & 0 \\ -e^{t_1} & e^{t_2} & 0 & \dots & 0 \\ 0 & -e^{t_2} & e^{t_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & e^{t_n} \end{pmatrix} \begin{pmatrix} X_{t_1} \\ X_{t_2} \\ X_{t_3} \\ \vdots \\ X_{t_n} \end{pmatrix}$$

Несложно понять, что матрица перехода служит матрицей Якоби и её определитель равен $e^{t_1+\dots+t_n}$. Следовательно,

$$p_{\boldsymbol{\eta}}(x_1, \dots, x_n) = e^{t_1+\dots+t_n} p_{\boldsymbol{\xi}}(e^{t_1}x_1, e^{t_2}x_2 - e^{t_1}x_1, \dots, e^{t_n}x_n - e^{t_{n-1}}x_{n-1}).$$

Подставляя, получаем, что

$$p_{\boldsymbol{\eta}}(x_1, \dots, x_n) = \frac{e^{t_1}}{\sqrt{2\pi e^{t_1}}} \exp\left\{-\frac{x_1^2 e^{2t_1}}{2e^{2t_1}}\right\} \prod_{k=2}^{n} \frac{e^{t_k}}{\sqrt{2\pi(e^{2t_k} - e^{2t_{k-1}})}} \exp\left\{-\frac{(e^{t_k}x_k - e^{t_{k-1}}x_{k-1})^2}{2(e^{2t_k} - e^{2t_{k-1}})}\right\} =$$
$$= \left(2\pi \prod_{k=2}^{n} \left(1 - e^{2(t_{k-1}-t_k)}\right)\right)^{-\frac{1}{2}} \exp\left\{-\frac{x_1^2}{2} - \frac{1}{2}\sum_{k=2}^{n} \frac{(x_k - x_{k-1}e^{t_{k-1}-t_k})^2}{1 - e^{2(t_{k-1}-t_k)}}\right\}.$$

1.2.3 Пуассновский процесс

Перейдём к одному из самых простых для исследования процессов — к *пуассоновскому потоку*.

Определение 7. Случайный процесс $X = (X_t)_{t \geqslant 0}$ называется пуассоновским потоком с *интенсивностью* λ, если он удовлетворяет трём условиям:

1. $X_0 = 0$ почти наверное.
2. X — процесс с независимыми приращениями.
3. Для всех $0 \leqslant s < t < +\infty$ $X_t - X_s \sim \mathrm{Pois}(\lambda(t-s))$, то есть для любого $n \in \mathbb{Z}_+$

$$\mathsf{P}(X_t - X_s = n) = \frac{(\lambda(t-s))^n}{n!} e^{\lambda(t-s)}.$$

Теперь построим пример такого процесса. Для этого вспомним процесс восстановления, описанный в примере **??**.

Пусть для любого натурального n T_n — это iid случайные величины с распределением $\mathrm{Exp}(\lambda)$, то есть их плотность равна $p(z) = \lambda e^{-\lambda z}\, \mathrm{I}\{z \geqslant 0\}$. Дальше, $S_n = T_1 + \ldots + T_n$, а для любого $t \geqslant 0$

$$N_t = \sum_{n=1}^{\infty} \mathrm{I}\{S_n \leqslant t\} = \#\{n \in \mathbb{N} : S_n < t\}.$$

Полученный случайный процесс обладает некоторыми интересными свойствами.

Свойство 4. *Для любого* $t > 0$ $N_t \sim \mathrm{Pois}(\lambda t)$.

Доказательство. Понятно, что N_t принимает значения в $\mathbb{Z}_+$. Следовательно, достаточно показать, что вероятности принять нужное значение будут именно такими, какими они должны быть.

Пусть $n = 0$. Тогда

$$\mathsf{P}(N_t = 0) = \mathsf{P}(T_1 > t) = \int\limits_t^\infty \lambda e^{-\lambda x}\,dx = e^{-\lambda t}\int\limits_t^\infty \lambda e^{-\lambda(x-t)}\,dx = e^{-\lambda t} = \frac{(\lambda t)^0}{0!}e^{-\lambda t}.$$

Теперь посмотрим на вероятность события $N_t = n$. Она равна

$$\mathsf{P}(N_t = n) = \mathsf{P}(N_t \geqslant n) - \mathsf{P}(N_t \geqslant n+1) = \mathsf{P}(S_n \leqslant t) - \mathsf{P}(S_{n+1} \leqslant t).$$

Для того, чтобы посчитать полученные вероятности, вспомним один факт: сумма n iid случайных величин с распределением $\mathrm{Exp}(\lambda)$ имеет гамма-распределение $\Gamma(1/\lambda, n)$ с плотностью

$$p_n(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}.$$

Тогда эта вероятность равна

$$\mathsf{P}(N_t = n) = \int\limits_0^t \left(\frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} - \frac{\lambda^{n+1} x^n e^{-\lambda x}}{n!}\right) dx = \int\limits_0^t \left(\frac{\lambda^n x^n e^{-\lambda x}}{n!}\right)' dx = \frac{(\lambda t)^n}{n!} e^{-\lambda t}. \qquad \square$$

Следствие. $\mathsf{E}[N_t] = \lambda t$, $\mathsf{D}[N_t] = \lambda t$.

Свойство 5. $N = (N_t)_{t\geqslant 0}$ *— это пуассоновский поток с интенсивностью* λ.

Доказательство. Для начала покажем, что $N_0 = 0$ почти наверное. Действительно, если $N_0 \neq 0$, то $T_1 = 0$, что происходит с нулевой вероятностью. Тем самым $N_0 = 0$ почти наверное.

Теперь докажем, что $(N_t)_{t\geqslant 0}$ удовлетворяет двум последним свойствам. Для этого заметим, что

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ \vdots \\ T_n \end{pmatrix}$$

Так как у матрицы единичный определитель и замена линейна, то

$$p_{S_1,\dots,S_n}(x_1, x_2, \dots, x_n) = p_{T_1,\dots,T_n}(x_1, x_2 - x_1, \dots, x_n - x_{n-1})$$

Пользуясь независимостью T_n, получаем, что

$$p_{S_1,\dots,S_n}(x_1, x_2, \dots, x_n) = \lambda e^{-\lambda x_1} \prod_{k=2}^n \lambda e^{-\lambda(x_k - x_{k-1})}\,\mathrm{I}\{x_k - x_{k-1} \geqslant 0\}$$

После преобразования получаем, что

$$p_{S_1,\dots,S_n}(x_1, x_2, \dots, x_n) = \lambda^n e^{-\lambda x_n}\,\mathrm{I}\{x_n \geqslant x_{n-1} \geqslant \ldots \geqslant x_1 \geqslant 0\}$$

Дальше, зафиксируем какие-либо числа $0 \leqslant t_1 < t_2 < \dots < t_n$, $k_1 \leqslant k_2 \leqslant \ldots \leqslant k_n$ и посмотрим на следующую вероятность:

$$\mathsf{P}\big(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2 - k_1, \dots, N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}\big)$$

Поймём, как связать это с S_n. Для этого поймём, как устроено первое условие. Оно означает, что $S_1, \ldots, S_{k_1} \leqslant t_1$, а $S_{k_1+1} > t_1$. Аналогично, получаем, что эта вероятность равна

$$\mathsf{P}\big(S_1, \ldots, S_{k_1} \in (0, t_1], S_{k_1+1}, \ldots, S_{k_2} \in (t_1, t_2], \ldots, S_{k_{n-1}+1}, \ldots, S_{k_n} \in (t_{n-1}, t_n], S_{k_n+1} > t_n\big)$$

Пользуясь плотностью случайного вектора из S_k, запишем это в виде интеграла:

$$\int\limits_{\substack{0<x_1,\ldots,x_{k_1}\leqslant t_1\\ t_1<x_{k_1+1},\ldots,x_{k_2}\leqslant t_2\\ \ldots\\ t_{n-1}<x_{k_{n-1}+1},\ldots,x_{k_n}\leqslant t_n\\ x_{k_n+1}>t_n}} \cdots \int \lambda^{k_n+1} e^{-\lambda x_{k_n+1}}\, \mathrm{I}\{x_{k_n+1} \geqslant x_{k_n} \geqslant \ldots \geqslant x_1 \geqslant 0\}\, dx_1 \ldots dx_{k_n}\, dx_{k_n+1}$$

Разобъём его в произведение интегралов, положив $k_0 = t_0 = 0$:

$$\lambda^{k_n} \int\limits_{t_n}^{\infty} \lambda e^{-\lambda x_{k_n+1}}\, dx_{k_n+1} \prod_{j=1}^{n} \int\limits_{t_{j-1}<x_{k_{j-1}+1}\leqslant\ldots\leqslant x_{k_j}\leqslant t_j} dx_{k_{j-1}+1} \ldots dx_{k_j}$$

Интегралы в произведении берутся достаточно просто: это объём симплекса. Рассуждая по аналогии с трёхмерным случаем, получаем, что интеграл равен

$$\lambda^{k_n} e^{-\lambda t_n} \prod_{j=1}^{n} \frac{(t_j - t_{j-1})^{k_j - k_{j-1}}}{(k_j - k_{j-1})!}$$

Теперь сделаем подгон: заметим, что

$$k_n = k_n - k_0 = \sum_{j=1}^{n} (k_j - k_{j-1}), \quad t_n = t_n - t_0 = \sum_{j=1}^{n} (t_j - t_{j-1})$$

Тогда интеграл равен

$$\prod_{j=1}^{n} \frac{(\lambda(t_j - t_{j-1}))^{k_j - k_{j-1}}}{(k_j - k_{j-1})!} e^{-\lambda(t_j - t_{j-1})}.$$

Какая красота. Отсюда мы сразу получаем оба свойства. Тем самым процесс восстановления для экспоненциального распределения является пуассоновским потоком с интенсивностью λ. □

Следующее свойство связано с понятием стационарных приращений.

Определение 8. Случайный процесс $X = (X_t)_{t\geqslant 0}$ имеет *стационарные приращения*, если для любых $0 \leqslant t_0 < t_1 < t_2 < \ldots < t_n < +\infty$ и $\forall h \geqslant 0$

$$(X_{t_1+h} - X_{t_0+h}, \ldots, X_{t_n+h} - X_{t_{n-1}+h}) \overset{d}{=} (X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}).$$

Свойство 6. *Пуассоновский поток имеет стационарные приращения.*

Доказательство. Следует из того, что $X_{a+h} - X_{b+h} \sim \mathrm{Pois}(a-b) \sim X_a - X_b$ и приращения независимы. □

Свойство 7. *Пусть $N^1, \ldots, N^k$ — независимые пуассоновские потоки с интенсивностями $\lambda_1, \ldots, \lambda_k$. Тогда случайный процесс $N_t = N_t^1 + \ldots + N_t^k$ — это пуассоновский процесс с интенсивностью $\lambda = \lambda_1 + \ldots + \lambda_k$.*

1.3 Стационарность случайных процессов

На практике часто попадаются процессы, которые неизменны во времени. Их принято называть *стационарными*. Это понятие было введено и для случайных процессов, хоть и не в одной ипостаси.

Буквальный перевод вышесказанного на математический язык даёт *стационарность в узком смысле*.

Определение 9. Случайный процесс $X = (X_t)_{t \in T}$ называется сильно стационарным (strong sense stationary, SSS, стационарным в узком смысле), если для любого натурального n, любых индексов $t_1, t_2, \ldots, t_n \in T$ и любого сдвига $h \geqslant 0$

$$F_{t_1+h,\ldots,t_n+h}(x_1,\ldots,x_n) = F_{t_1,\ldots,t_n}(x_1,\ldots,x_n).$$

Пример 5. Последовательность iid случайных величин — это стационарная в узком смысле случайная последовательность.

Пример 6. Винеровский процесс не является стационарным в узком смысле. Действительно, пусть $n = 1$ и $h > 0$. Тогда $B_t \sim \mathcal{N}(0, t)$, а $B_{t+h} \sim \mathcal{N}(0, t+h)$ и $F_{t+h}(x) \neq F_t(x)$.

У стационарных в узком смысле процессов есть одно полезное свойство. Но у него есть требование — процесс должен быть второго порядка. Что это значит?

Определение 10. Случайный процесс $(X_t)_{t \in T}$ называется *процессом второго порядка*, если $\mathsf{E}[X_t^2]$ конечно для всех t (то есть это ограниченная функция от t).

Свойство 1. *Если $(X_t)_{t \in T}$ — стационарный в узком смысле процесс второго порядка, то*

1. $m(t) = \mu = const$, $D(t) = \sigma^2 = const$.
2. *Для любых* $t_1, t_2, h \in T$ $R_X(t_1, t_2) = R_X(t_1 + h, t_2 + h)$.

Данное свойство очевидным образом следует из определения стационарного в узком смысле процесса и того, что матожидание, дисперсия и ковариация конечны. Последнее свойство позволяет свести ковариационную функцию к одному аргументу: $R_X(t_1, t_2) = R_X(0, t_2 - t_1) \equiv R_X(t_2 - t_1)$.

Вообще говоря, выполнение этих трёх свойств — это тоже в некоторой степени стационарность. Только в широком смысле.

Определение 11. Случайный процесс $X = (X_t)_{t \in T}$ называется слабо стационарным (стационарным в широком смысле, wide sense stationary, WSS, ковариационно стационарным, стационарным второго порядка),[2] если выполнены следующие условия:

1. $m(t) = \mu = const$, $D(t) = \sigma^2 = const$.
2. Для любых $t_1, t_2, h \in T$ $R_X(t_1, t_2) = R_X(t_1 + h, t_2 + h)$.

Обычно сильная и слабая стационарности идут вместе (если это не так, то что-то пошло не туда). Например, если семейство конечномерных функций распределения полностью задаётся матожиданием и ковариационной функцией, то сильная стационарность равносильна слабой стационарности.

У ковариационной функции стационарного в широком смысле случайного процесса есть несколько свойств:

[2] Название ковариационной стационарности появилось из-за того, что ковариационная функция зависит только от разности индексов. Стационарность второго порядка же означает постоянство второго момента.

1. Она неотрицательна в нуле: $R_X(0) = R_X(t,t) = D(t) = \sigma^2 \geqslant 0$.

2. Она чётна: $R_X(-\tau) = R_X(0,-\tau) = R_X(-\tau,0) = R_X(0,\tau) = R_X(\tau)$.

3. Она ограничена по модулю дисперсией. Действительно, по неравенству Коши-Буняковского-Шварца $|R_X(\tau)| = |R_X(t,t+\tau)| = |\operatorname{\mathsf{cov}}(X_t, X_{t+\tau})| \leqslant \sqrt{\mathsf{D}[X_t]\,\mathsf{D}[X_{t+\tau}]} = \sigma^2$.

4. Аналог неотрицательной определённости: для любого натурального n, любого неслучайного вектора $(z_1,\ldots,z_n)$ и любого набора индексов $t_1,\ldots,t_n$

$$\sum_{i,j=1}^{n} R_X(t_i - t_j) z_i z_j \geqslant 0.$$

5. Если $R_X(\tau)$ непрерывна в нуле, то она непрерывна для любого τ.

Доказательство. Пусть $\xi_t = X_t - \mathsf{E}[X_t]$. Пользуясь этим обозначением, распишем разность ковариационных функций.

$$R_X(t_1+h_1, t_2+h_2) - R_X(t_1,t_2) = \mathsf{E}[\xi_{t_1+h_1}\xi_{t_2+h_2} - \xi_{t_1}\xi_{t_2}] = \\ = \mathsf{E}[\xi_{t_1+h_1}(\xi_{t_2+h_2} - \xi_{t_2})] + \mathsf{E}[(\xi_{t_1+h_1} - \xi_{t_1})\xi_{t_2}]$$

Далее, по неравенству треугольника:

$$|R_X(t_1+h_1, t_2+h_2) - R_X(t_1,t_2)| \leqslant |\,\mathsf{E}[\xi_{t_1+h_1}(\xi_{t_2+h_2} - \xi_{t_2})]| + |\,\mathsf{E}[(\xi_{t_1+h_1} - \xi_{t_1})\xi_{t_2}]|$$

Теперь воспользуемся неравенством Коши-Буняковского-Шварца:

$$|R_X(t_1+h_1, t_2+h_2) - R_X(t_1,t_2)| \leqslant \sqrt{\mathsf{E}\left[\xi^2_{t_1+h_1}\right]\mathsf{E}[(\xi_{t_2+h_2} - \xi_{t_2})^2]} + \sqrt{\mathsf{E}[(\xi_{t_1+h_1} - \xi_{t_1})^2]\,\mathsf{E}\left[\xi^2_{t_2}\right]}$$

Осталось показать, что эта сумма стремится к нулю. Покажем, что первый член уходит в ноль (второй рассматривается аналогично). Действительно, по непрерывности $R_X(t,t)$

$$\mathsf{E}\left[(\xi_{t_2+h_2} - \xi_{t_2})^2\right] = \mathsf{E}\left[\xi^2_{t_2+h_2}\right] - 2\,\mathsf{E}[\xi_{t_2+h_2}\xi_{t_2}] + \mathsf{E}\left[\xi^2_{t_2}\right] = R_X(t_2+h_2, t_2+h_2) + \\ + R_X(t_2,t_2) - 2R_X(t_2+h_2, t_2) \xrightarrow[h_2\to 0]{} 0.$$

Следовательно, $R_X(t_1+h_1, t_2+h_2) \xrightarrow[h_1,h_2\to 0]{} R_X(t_1,t_2)$ и $R_X(t_1,t_2)$ непрерывна везде. □

Теперь посмотрим на три класса случайных процессов: IID, SSS и WSS. Есть ли между ними какая-либо связь? Есть. Начнём с очевидной цепочки вложений: IID $\subseteq$ SSS $\subseteq$ WSS. Хотя второе вложение не совсем корректно — не все сильно стационарные процессы являются процессами второго порядка. Если добавить это требование, то вложение станет корректным. Теперь покажем, что все вложения строгие.

- Начнём с SSS $\setminus$ IID. Возьмём случайную последовательность $X = (X_t)_{t\in T}$, устроенную следующим образом: для всех t $X_t = \xi$, где ξ — это какая-то фиксированная случайная величина. Она очевидно является стационарной в сильном смысле и все её сечения одинаково распределены, но она не задаёт последовательность независимых случайных величин.

- Теперь посмотрим на WSS $\setminus$ SSS. Суть примера в том, что мы будем брать разные распределения, у которых совпадают матожидание и дисперсия. Например, пусть

$X = (X_t)_{t\geqslant 0}$ — случайная последовательность независимых случайных величин такая, что $X_{2n} \sim \mathrm{Bern}(p)$, а $X_{2n+1} \sim \mathcal{N}(p, p(1-p))$. Понятно, что ни о каком равенстве распределений и речи быть не может, а вот слабая стационарность выполнена (почему?).

- Приведём ещё один пример процесса из SSS $\setminus$ IID. Пусть $\{X_n\}_{n=1}^{\infty}$ — последовательность iid случайных величин. Построим по ней новую последовательность $\{Y_n\}_{n=1}^{\infty}$ по правилу $Y_n = X_n + X_{n+1}$. Понятно, что полученная последовательность не будет состоять из независимых случайных величин, но она будет сильно стационарной.

2 Генерирование реализаций случайных процессов

Теперь посмотрим, как симулировать различные процессы. Начнём с самого простого — с пуассоновского потока.

2.1 Генерирование пуассоновских случайных процессов

2.1.1 Однородный пуассоновский поток событий

Для тех, кто забыл — определение пуассоновского потока дано в примере **??**. Единственная сложность в генерации реализации состоит в том, что нужно уметь генерирвать случайные величины из экспоненциального распределения. Но мы можем свободно генерировать случайные величины из $\mathrm{U}(0,1)$. Как получить из него экспоненциальное распределение? Для этого докажем одно утверждение:

Лемма (Метод обратного преобразования)**.** *Пусть X — случайная величина с неубывающей функцией распределения $F : \mathbb{R} \mapsto [0,1]$. Введём обратную функцию $F^{-1} : [0,1] \mapsto \mathbb{R}$ следующим образом: $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geqslant u\}$. Тогда, если $U \sim \mathrm{U}(0,1)$, то $F^{-1}(U)$ имеет функцию распределения F.*

Доказательство. Действительно, $\mathsf{P}(F^{-1}(U) \leqslant x) = \mathsf{P}(U \leqslant F(x)) = F(x)$. □

Теперь покажем, как генерировать случайные величины из распределения $\mathrm{Exp}(\lambda)$. Рассмотрим функцию распределения:

$$F(x) = 1 - e^{-\lambda x} \implies x = -\frac{1}{\lambda}\ln(1 - F(x)) \implies F^{-1}(x) = -\frac{1}{\lambda}\ln(1 - x).$$

Тогда по методу обратного преобразования $-\ln(1-U)/\lambda$ будет иметь распределение $\mathrm{Exp}(\lambda)$. Теперь заметим, что $1 - U \overset{d}{=} U$. Тогда получаем, что $-\frac{1}{\lambda}\ln U$ будет иметь нужное распределение.

Теперь несложно написать алгоритм генерации реализации однородного пуассоновского потока.

Алгоритм 1 Алгоритм генерации реализации однородного пуассоновского потока

Вход: Интенсивность λ, максимальное время T.

1: $t \leftarrow 0, I \leftarrow 0, S \leftarrow \varnothing$
2: сгенерировать $U \sim \mathrm{U}(0,1)$
3: $t \leftarrow t - \ln(U)/\lambda$

while $t \leqslant T$ **do**
    $I \leftarrow I + 1, S(I) \leftarrow t$
    сгенерировать $U \sim \mathrm{U}(0, 1)$
    $t \leftarrow t - \ln(U)/\lambda$

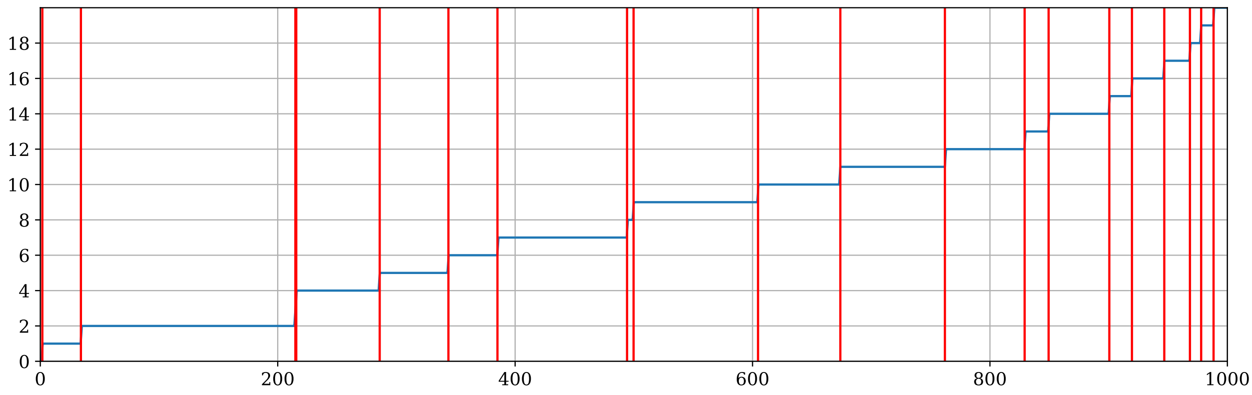

Рис. 1: Пример реализации пуассоновского потока с параметрами $T = 1000$, $\lambda = 0.02$

2.1.2 Неоднородный пуассоновский поток событий

Ранее мы смотрели на однородный пуассоновский процесс. Однородный он по той простой причине, что его интенсивность постоянна. Теперь скажем, что λ — это какая-то функция от t. В таком случае получим *неоднородный пуассоновский поток*. Определяется он почти так же, как и однородный, только немного изменяется третье свойство:

$$X_t - X_s \sim \mathrm{Pois}\left(\int\limits_s^t \lambda(x)\,dx\right).$$

Но считать интегралы не очень приятно. Можно ли обойтись без них? Можно. Рассмотрим однородный пуассоновский поток N_t с интенсивностью λ. Пусть событие, появляющееся в момент времени t “засчитывается” с некоторой вероятностью $p(t)$, то есть

$$\mathsf{P}(N_t = N_{t-\varepsilon} + 1) = p(t), \quad \mathsf{P}(N_t = N_{t-\varepsilon}) = 1 - p(t)$$

Оказывается, что N_t — неоднородный пуассоновский поток событий с интенсивностью $\lambda(t) = \lambda p(t)$. Этот результат называется *теоремой Льюиса-Шедлера*.

Пусть $\tilde{\lambda} = \max_{t \in [0,T]} \lambda(t)$. Тогда алгоритм будет выглядеть так:

Алгоритм 2 Алгоритм генерации реализации неоднородного пуассоновского потока

Вход: Интенсивность $\lambda(t)$, максимальное время T.
$t \leftarrow 0, I \leftarrow 0, S \leftarrow \varnothing,$
сгенерировать $U_1 \sim \mathrm{U}(0, 1)$
$t \leftarrow t - \ln(U_1)/\tilde{\lambda}$
while $t \leqslant T$ **do**
    сгенерировать $U_2 \sim \mathrm{U}(0, 1)$
    if $U_2 \leqslant \lambda(t)/\tilde{\lambda}$ **then**
        $I \leftarrow I + 1, S(I) \leftarrow t$
    сгенерировать $U_1 \sim \mathrm{U}(0, 1)$
    $t \leftarrow t - \ln(U_1)/\tilde{\lambda}$

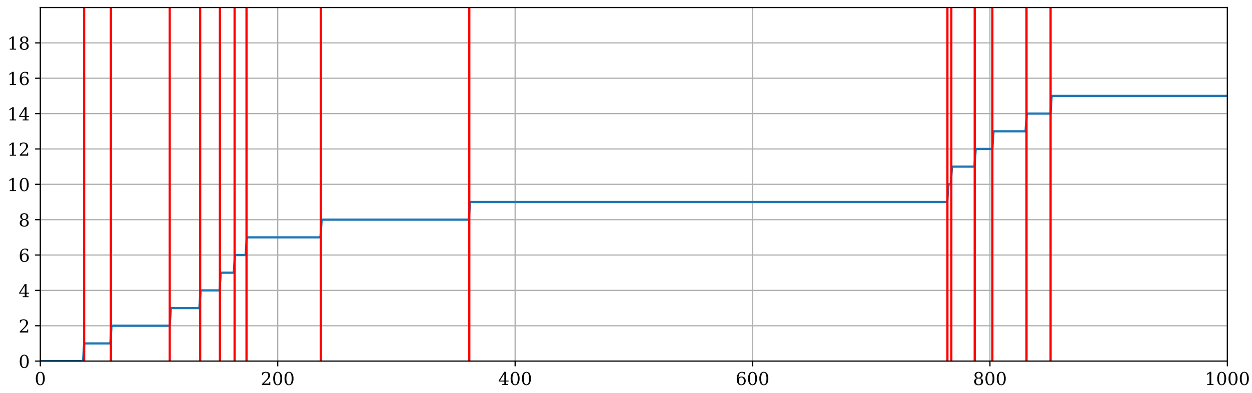

Рис. 2: Пример реализации неоднородного пуассоновского потока с параметрами $T = 1000$, $\lambda(t) = (\sin(t/100) + 1)/100$.

2.2 Метод стохастического интегрирования

Как известно, дифференциальные уравнения описывают очень многое. Но, оказывается, их можно приспособить и для описания случайных процессов. Основное отличие состоит в том, что в данном случае функция, относительно которой решается уравнение, является случайной величиной. Такие дифференциальные уравнения называют *стохастическими.*

Оказывается, что многие процессы, которые изучаются на практике, "управляются" броуновским движением. Однако есть проблема: траектории винеровского процесса нигде не дифференцируемы почти наверное. Поэтому манипулирование с процессами такого типа потребовало создания собственного исчисления, называемого теорией *стохастических интегралов.* Дадим определение:

Определение 12. Пусть $T = \{t_k\}_{k=0}^n$ — некоторое разбиение отрезка $[0, t]$. Далее, выберем точки $\tau = \{\tau_k\}_{k=1}^n$ по правилу $\tau_k \in [t_{k-1}, t_k]$. Составим по этому разбиению интегральную сумму:

$$S_n = \sum_{k=1}^{n} b(\tau_k)(B_{t_k} - B_{t_{k-1}})$$

Стохастическим интегралом от неслучайной функции $b(t)$ по броуновскому движению $B = (B_t)_{t\geqslant 0}$ называют предел интегральных сумм при диаметре разбиения, стремящемся к нулю:[3]

$$\int\limits_0^t b(x)\, dB_x = \lim_{\Delta T \to 0} S_n$$

Примечание. Не стоит забывать, что $B_{t_i} - B_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$. Это поможет при симуляции процесса.

Многие стохастические процессы могут быть записаны в виде

$$X_t = \int\limits_0^t a(x)\, dx + \int\limits_0^t b(x)\, dB_x,$$

[3]Вопрос о том, почему этот предел вообще существует и каким образом последовательность частичных сумм сходится к нему, оставим за кадром.

где $a(x)$ и $b(x)$ — некоторые неслучайные функции. Это же выражение можно записать в дифференциалах: $dX_t = a(t)\,dt + b(t)\,dB_t$.

Как использовать этот метод? Примерно так же, как и в численном интегрировании: заменить дифференциал на малое изменение и суммировать.

$$X_{t+\varepsilon} - X_t \approx a(t)\varepsilon + b(t)(B_{t+\varepsilon} - B_t).$$

2.3 Метод гауссовских векторов

Если нужно сгенерировать не слишком большую реализацию гауссовского процесса, то ситуация становится несколько проще. Как известно, у них все конечномерные функции распределения являются гауссовскими, поэтому реализация будет являться гауссовским вектором. Далее, нам известны математическое ожидание и ковариационная функция процесса. Из этого можно вытащить математическое ожидание и матрицу ковариаций нужного вектора.

Но есть проблема: как генерировать случайный гауссовский вектор с заданным распределением? Сходу это сделать не получится. Для этого проведём одно рассуждение.

Допустим, что вектор одномерный, то есть это просто случайная величина $\xi \sim \mathcal{N}(\mu, \sigma^2)$. Как из стандартного нормального распределения получить нужное распределение? Легко: $\xi \overset{d}{=} \mu + \sigma\eta$, где $\eta \sim \mathcal{N}(0,1)$.

Оказывается, что для общего случая верно нечто похожее. Пусть $\boldsymbol{\mu}$ — некоторый фиксированный вектор, $\mathbf{\Sigma}$ — квадратная матрица, а $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{E})$. Какое распределение будет у случайного вектора $\boldsymbol{\eta} = \boldsymbol{\mu} + \mathbf{\Sigma}\boldsymbol{\xi}$? Так как преобразование линейно, то это гауссовский вектор с параметрами

$$\begin{aligned} \mathsf{E}[\boldsymbol{\eta}] &= \mathsf{E}[\boldsymbol{\mu} + \mathbf{\Sigma}\boldsymbol{\xi}] = \mathsf{E}[\boldsymbol{\mu}] + \mathbf{\Sigma}\,\mathsf{E}[\boldsymbol{\xi}] = \boldsymbol{\mu} \\ \mathsf{D}[\boldsymbol{\eta}] &= \mathsf{E}[\mathbf{\Sigma}\boldsymbol{\xi}(\mathbf{\Sigma}\boldsymbol{\xi})^\intercal] = \mathbf{\Sigma}\,\mathsf{E}[\boldsymbol{\xi}\boldsymbol{\xi}^\intercal]\mathbf{\Sigma}^\intercal = \mathbf{\Sigma}\mathbf{\Sigma}^\intercal. \end{aligned}$$

Теперь вспомним один факт из линейной алгебры.

Теорема 6 (Разложение Холецкого)**.** *Пусть $\mathbf{\Sigma}$ — неотрицательно определённая симметричная матрица. Тогда существует нижнетреугольная матрица $\mathbf{C}$ с неотрицательными членами на диагонали такая, что $\mathbf{\Sigma} = \mathbf{C}\mathbf{C}^\intercal$. Если же $\mathbf{\Sigma}$ положительно определена, то все члены $\mathbf{C}$ на диагонали строго положительны.*

Выпишем формулы для вычисления матрицы $\mathbf{C}$:

$$\mathbf{C}_{kk} = \sqrt{\mathbf{\Sigma}_{kk} - \sum_{j=1}^{k-1} \mathbf{C}_{ki}^2}, \quad \mathbf{C}_{ij} = \frac{1}{\mathbf{C}_{jj}}\left(\mathbf{\Sigma}_{ij} - \sum_{k=1}^{j-1} \mathbf{C}_{ik}\mathbf{C}_{jk}\right)$$

Отсюда понятно, как генерировать гауссовский вектор с заданным распределением $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$. Для этого мы независимо генерируем n случайных величин с распределением $\mathcal{N}(0,1)$ (это можно сделать с помощью того же преобразования Бокса-Мюллера) и получаем гауссовский вектор $\boldsymbol{\xi}$. Далее, берём матрицу $\mathbf{C}$ из разложения Холецкого и строим новый вектор $\boldsymbol{\eta} = \boldsymbol{\mu} + \mathbf{C}\boldsymbol{\xi}$. Полученный вектор будет иметь нужное распределение.

2.4 Генерирование гауссовских случайных процессов

2.4.1 Винеровский процесс

Сначала разберёмся, как генерировать его с помощью гауссовских векторов. Для этого достаточно сгенерировать матрицу ковариаций и вектор матожиданий. Как известно, $m(t) = 0$, а $R(s,t) = \min(s,t)$.

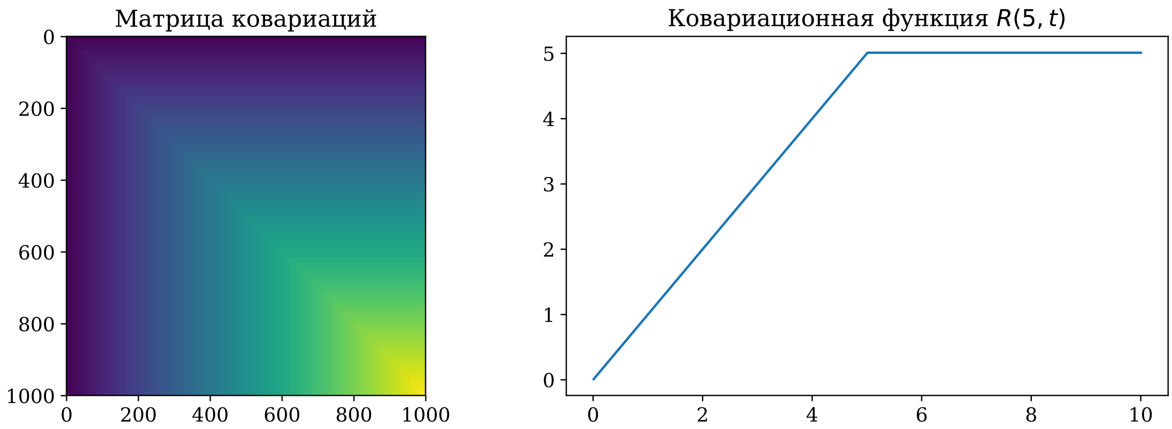

Рис. 3: Визуализация матрицы ковариаций и $R(5, t)$ для $t \in (0, 10)$.

Интереснее генерация с помощью стохастического интегрирования. Хотя и данном случае всё достаточно очевидно: B_t можно приблизить суммой достаточно большого числа номальных случайных величин:

$$B_t = \int_0^t dB_x \approx \sum_{k=1}^{N} (B_{t_k} - B_{t_{k-1}}).$$

Для примера посмотрим на первые десять секунд. Для этого разобьём отрезок $[0, 10]$ на 1000 равных кусков и посчитаем эту сумму. Это даст приемлемую точность.

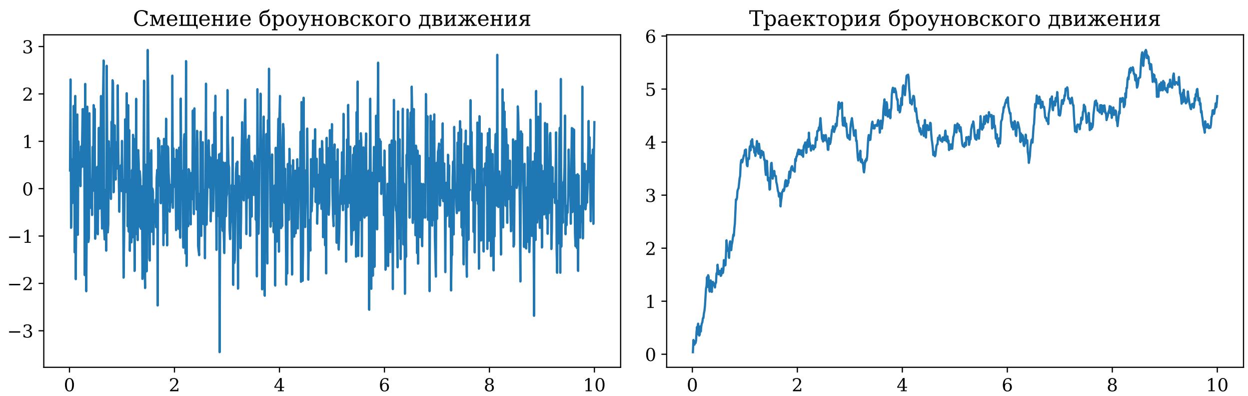

Рис. 4: Пример реализации первых десяти секунд броуновского движения.

2.4.2 Процесс Орнштейна-Уленбека

Ранее мы обсуждали процесс Орнштейна-Уленбека. Однако на самом деле он определяется немного по-другому:

Определение 13. Случайный процесс $X = (X_t)_{t \geqslant 0}$, удовлетворяющий стохастическому дифференциальному уравнению

$$dX_t = \theta(\mu - X_t)\, dt + \sigma\, dB_t, \quad X_0 = x_0,$$

называется процессом Орнштейна-Уленбека.

Для того, чтобы получить ранее описанный процесс, нужно подставить $\sigma = \sqrt{2}$, $\theta = 1$, $\mu = 0$ и $x_0 = 0$. Для генерации его реализации с помощью гауссовского вектора достаточно вспомнить, что $\mathsf{E}[X_t] = 0$ и $\mathsf{cov}(X_t, X_s) = e^{-|t-s|}$.

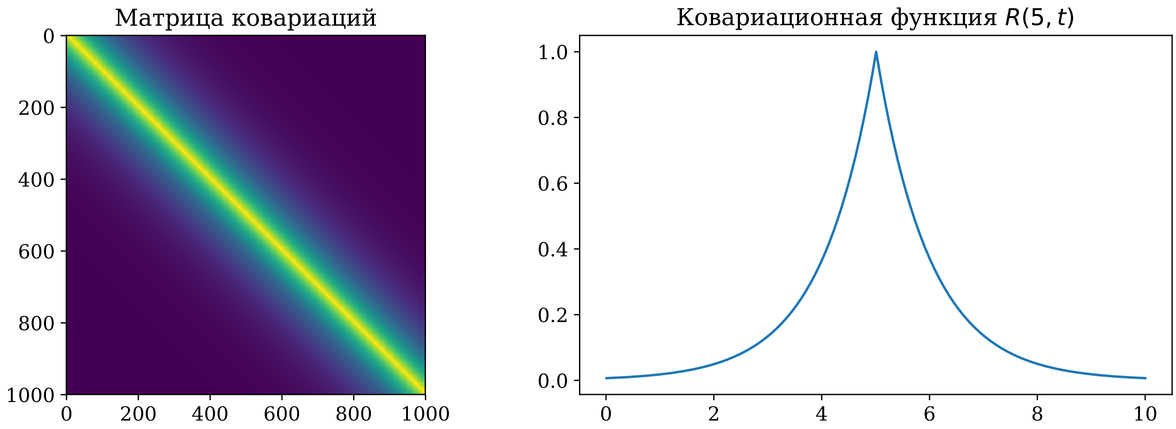

Рис. 5: Визуализация матрицы ковариаций и $R(5,t)$ для $t \in (0, 10)$.

Разностная схема устроена следующим образом:

$$X_{t_{i+1}} - X_{t_i} = \theta(\mu - X_{t_i})(t_{i+1} - t_i) + \sigma(B_{t_{i+1}} - B_{t_i}).$$

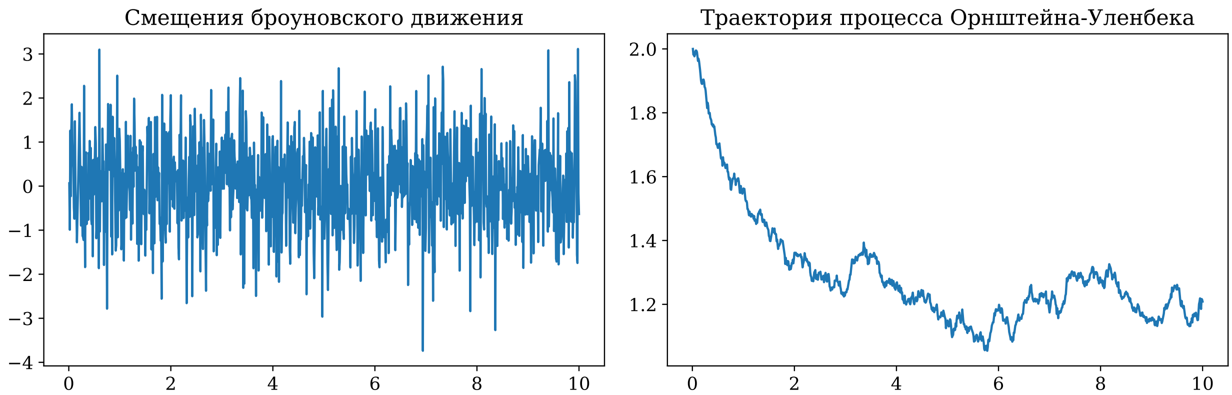

Рис. 6: Пример реализации первых десяти секунд процесса Орнштейна-Уленбека с параметрами $x_0 = 2$, $\sigma = 0.1$, $\mu = 1.2$, $\theta = 1$.

2.4.3 Фрактальное броуновское движение

Напоследок рассмотрим ещё один случайный процесс, называемый *фрактальным броуновским движением*.

Определение 14. Фрактальное броуновское движение с *параметром Хёрста* $H \in (0, 1)$ — это гауссовский случайный процесс с непрерывным временем $B^H = (B^H_t)_{t \in [0,T]}$, удовлетворяющий следующим условиям:

- $B^H_0 = 0$ почти наверное,
- $\mathsf{E}\big[B^H_t\big] = 0$ для всех $t \in [0, T]$,
- $\mathsf{cov}(B^H_t, B^H_s) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$.

Оказывается, что если подставить $H = 1/2$, то получится обычное броуновское движение.[4] В остальных случаях получается некоторый гауссовский процесс.

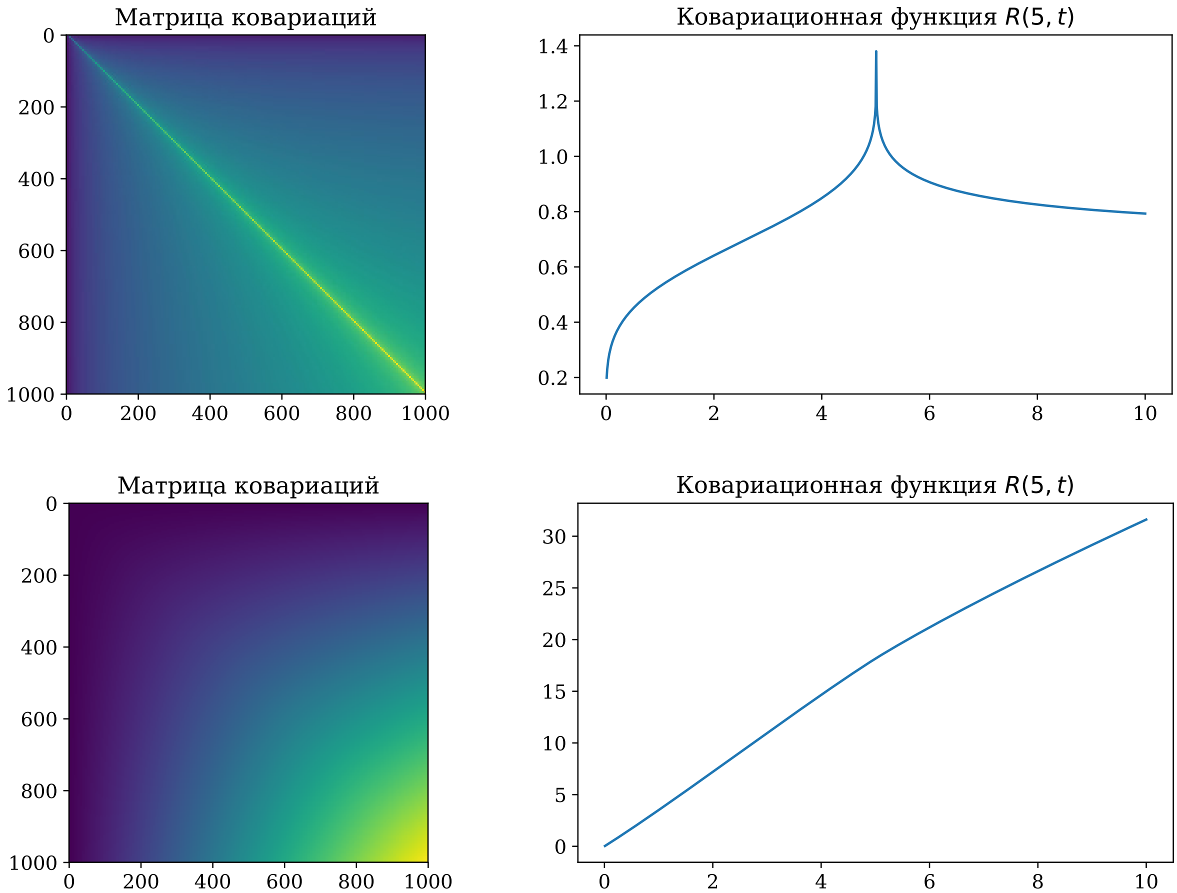

Рис. 7: Визуализация матрицы ковариаций и $R(5, t)$ для $t \in (0, 10)$ при $H = 0.1$ и $H = 0.9$.

После того, как была получена матрица ковариаций, дело остаётся за малым: получить нужнул реализацию с помощью разложения Холецкого. Я не буду описывать технические детали, ибо они и так очевидны. Теперь посмотрим, как себя ведут траектории в зависимости от параметра Хёрста.

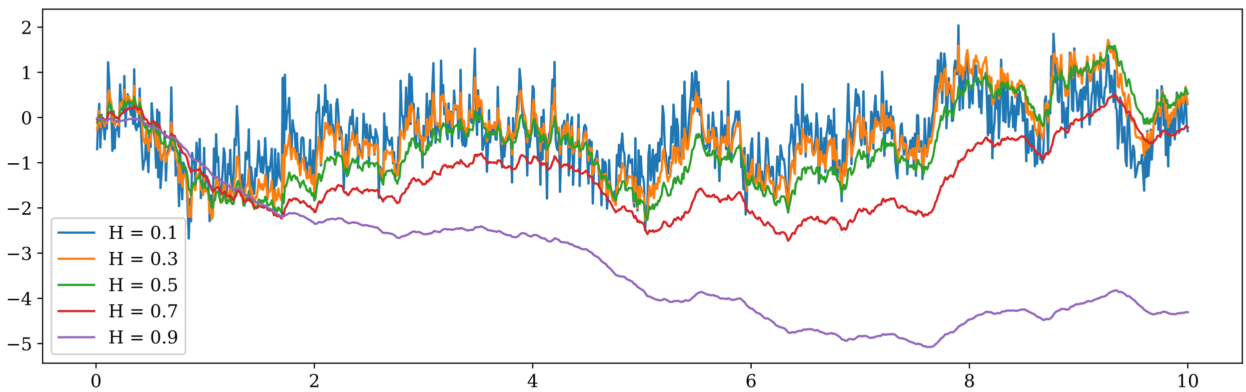

Рис. 8: Пример реализаций первых десяти секунд фрактального броуновского движения при разных параметрах Хёрста.

[4]Возникает вопрос о том, что делать с независимостью приращений. Но есть теорема, которая гласит, что фрактальное броуновское движение имеет независимые приращения только при $H = 1/2$.