

Неделя 3. Многомерное нормальное распределение и гауссовские процессы

Carlos Buitrago

Table of contents

1	1. Гауссовский случайный вектор	1
1.1	Симуляция 1. Корреляция меняет геометрию нормального облака	2
1.2	Важная ловушка: нормальные маргиналы не гарантируют совместную нормальность	2
2	2. Гауссовский процесс	3
3	3. Ковариационная функция и критерий существования	3
3.1	Быстрые проверки кандидата R	3
4	4. Как симулировать гауссовский процесс на сетке	4
5	5. Линейные преобразования процессов	4
6	6. Задача: какие преобразования гауссовы?	5
6.1	Симуляция 2. Почему возведение в квадрат ломает нормальность	5
7	7. Задача: какие функции могут быть ковариациями?	6
7.1	Симуляция 4. Постоянная ковариация $R(s, t) = 2$	6
8	8. Дополнительные примеры и связь с приложениями	7
8.1	Броуновское движение	7
8.2	Стационарные гауссовские процессы	7
8.3	Прогнозирование: условное распределение гауссовского вектора	7
8.4	Симуляция 5. GP-регрессия: прогноз и неопределённость	8

1 1. Гауссовский случайный вектор

i Определение

Вектор $\xi = (\xi_1, \dots, \xi_n)^\top$ называется **гауссовским**, если его характеристическая функция имеет вид

$$\varphi_\xi(u) = \mathbb{E} e^{iu^\top \xi} = \exp \left(iu^\top a - \frac{1}{2} u^\top \Sigma u \right), \quad u \in \mathbb{R}^n,$$

где $a \in \mathbb{R}^n$, а Σ — симметричная неотрицательно определённая матрица. Обозначение: $\xi \sim \mathcal{N}(a, \Sigma)$.

При этом $a = \mathbb{E}\xi$, $\Sigma_{ij} = \text{cov}(\xi_i, \xi_j)$. **Вырожденные** нормальные распределения тоже допускаются: например, константа c имеет закон $\mathcal{N}(c, 0)$.

! Три эквивалентные характеристики

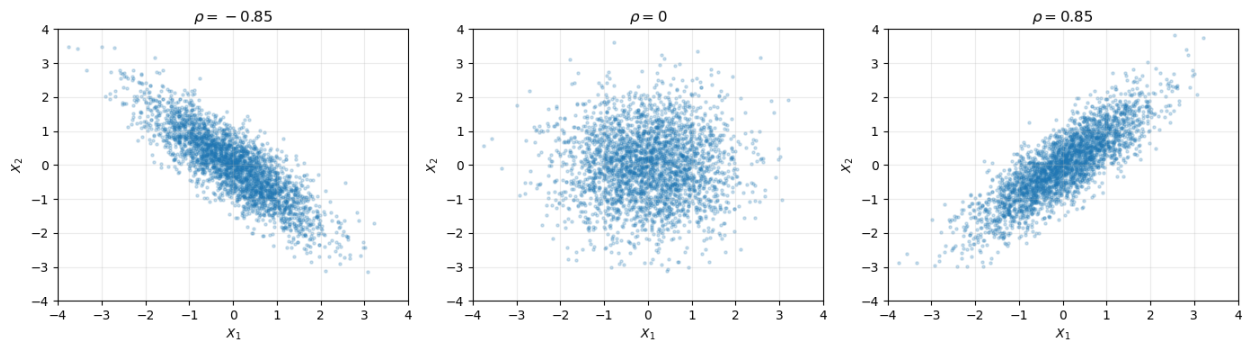
1. ξ — гауссовский вектор.
2. Существуют матрица $A \in \mathbb{R}^{n \times m}$, вектор $b \in \mathbb{R}^n$ и независимые $Z_j \sim \mathcal{N}(0, 1)$ такие, что $\xi = AZ + b$.
3. Для любого $u \in \mathbb{R}^n$ величина $u^\top \xi$ нормальна (включая вырожденный случай).

Почему это полезно? Чтобы доказать гауссовость нового вектора, достаточно представить его как детерминированное линейное преобразование исходного гауссовского вектора.

Независимость: если вектор совместно гауссовский, то его координаты независимы тогда и только тогда, когда попарно некоррелированы (Σ диагональна). Без совместной гауссовости это неверно.

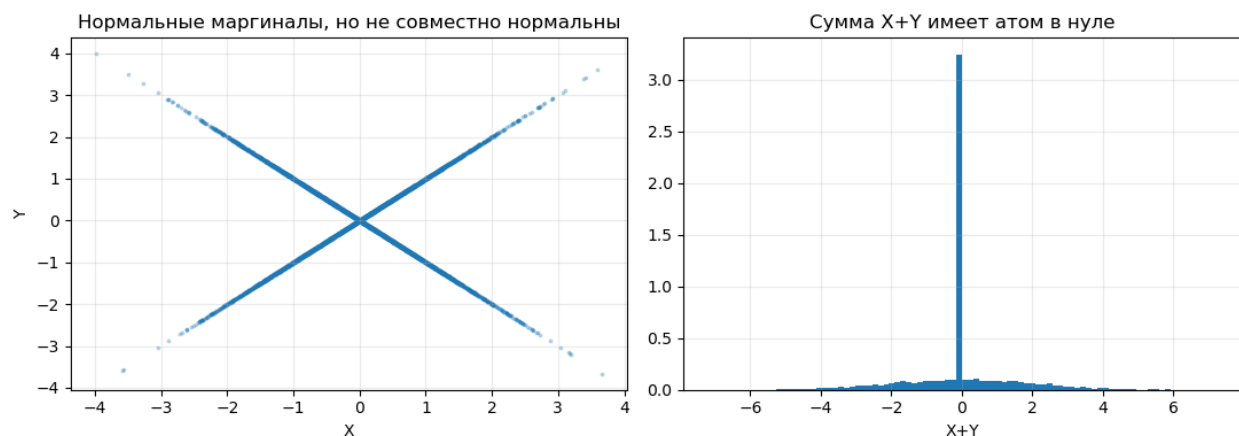
1.1 Симуляция 1. Корреляция меняет геометрию нормального облака

Сгенерируем $Z \sim \mathcal{N}(0, I_2)$ и положим $X = AZ$. Тогда $\text{cov}(X) = AA^\top$.



1.2 Важная ловушка: нормальные маргиналы не гарантируют совместную нормальность

Пусть $Z \sim \mathcal{N}(0, 1)$, S независимо от Z и принимает значения ± 1 с вероятностью $1/2$. Тогда $X = Z$ и $Y = SZ$ по отдельности стандартно нормальны и $\text{cov}(X, Y) = 0$, но $X + Y$ не нормально: с вероятностью $1/2$ оно равно нулю.



Эмпирическая ковариация: -0.010901520049899338

2. Гауссовский процесс

Определение

Процесс $(X_t)_{t \in T}$ называется **гауссовским**, если для любого $n \geq 1$ и любых $t_1, \dots, t_n \in T$ вектор

$$(X_{t_1}, \dots, X_{t_n})$$

совместно гауссовский.

Важно: мало проверить, что каждый X_t по отдельности нормально распределён. Нужно проверить *совместную* нормальность любого конечного набора времён.

Гауссовский процесс задаётся в смысле конечномерных распределений двумя функциями:

$$m(t) = \mathbb{E}X_t, \quad R(s, t) = \text{cov}(X_s, X_t).$$

Для $t_1, \dots, t_n$ имеем

$$(X_{t_1}, \dots, X_{t_n})^\top \sim \mathcal{N}(m_T, K_T), \quad (K_T)_{ij} = R(t_i, t_j).$$

Траектории и распределения: ковариационная функция определяет конечномерные законы, но сама по себе не означает непрерывность траекторий. Например, для независимых стандартных нормальных X_t при разных t ковариация равна $\mathbf{1}_{\{s=t\}}$; на непрерывном множестве индексов это не даёт непрерывных траекторий.

3. Ковариационная функция и критерий существования

Необходимое и достаточное условие

Функция $R : T \times T \rightarrow \mathbb{R}$ может быть ковариацией гауссовского процесса тогда и только тогда, когда она

1. **симметрична:** $R(s, t) = R(t, s)$;
2. **неотрицательно определена:** для всех $n, t_1, \dots, t_n, c_1, \dots, c_n$,

$$\sum_{i,j=1}^n c_i c_j R(t_i, t_j) \geq 0.$$

Функция среднего $m : T \rightarrow \mathbb{R}$ может быть произвольной.

Доказательство необходимости:

$$\sum_{i,j} c_i c_j R(t_i, t_j) = \text{Var} \left(\sum_i c_i X_{t_i} \right) \geq 0.$$

Симметрия следует из симметрии ковариации.

Идея достаточности: каждой конечной совокупности времён сопоставим многомерный нормальный закон с вектором средних $(m(t_i))$ и матрицей $(R(t_i, t_j))$. Положительная полуопределённость гарантирует существование этих законов, а их согласованность позволяет применить теорему Колмогорова о продолжении.

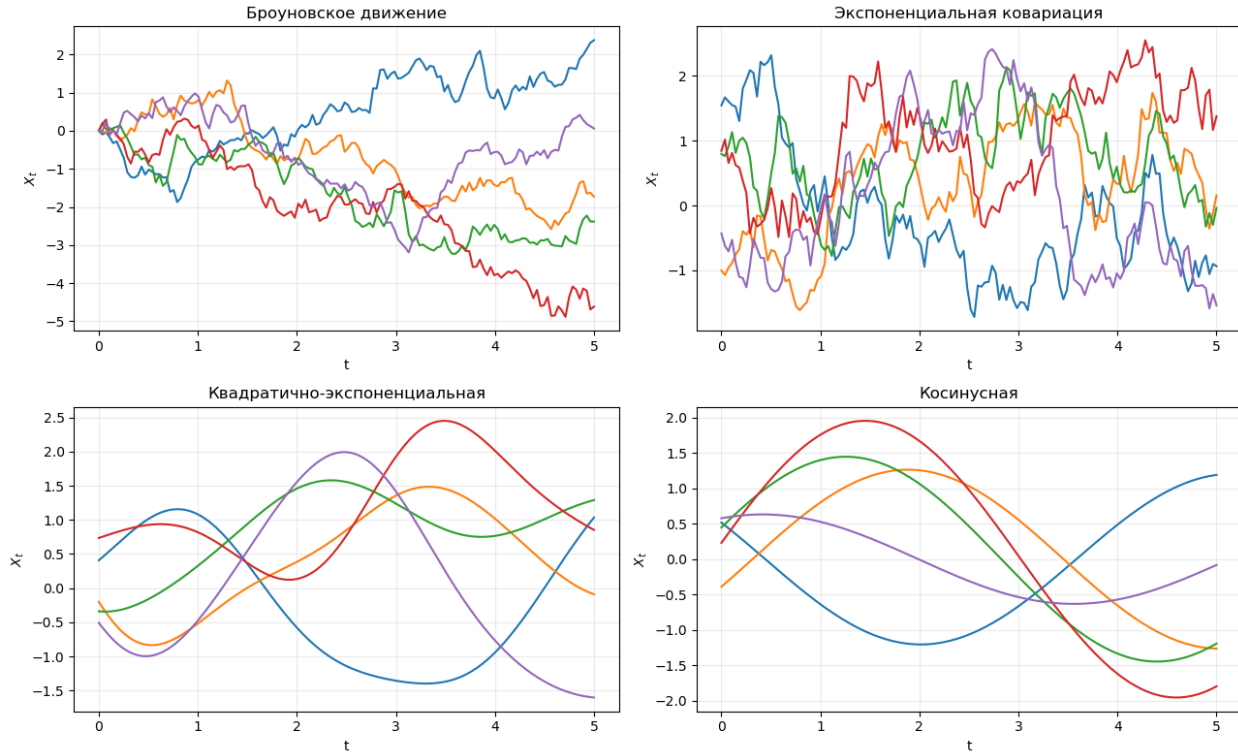
3.1 Быстрые проверки кандидата R

- **Диагональ:** $R(t, t) = \text{Var}(X_t) \geq 0$.
- **Неравенство Коши — Буняковского:** $|R(s, t)|^2 \leq R(s, s)R(t, t)$.
- **Матрицы Грама:** для любых выбранных времён все собственные значения $(R(t_i, t_j))$ неотрицательны.

Отрицательная диагональ или отрицательный определитель 2×2 немедленно исключают функцию. Однако проверка лишь нескольких сеток не является строгим доказательством положительной определённости для всех времён.

4. Как симулировать гауссовский процесс на сетке

Для $t_1, \dots, t_n$ строим $K = (R(t_i, t_j))$, затем генерируем $Z \sim N(0, I_n)$ и $X = m + LZ$, где $LL^\top = K$. Если K вырождена, используем спектральное разложение с обрезкой малых отрицательных ошибок округления. На графике изображается только дискретизация процесса; соединение точек линиями не доказывает непрерывность его траекторий.



5. Линейные преобразования процессов

! Основной принцип

Если (X_t) — гауссовский процесс, а f, g, h — **детерминированные** функции (причём $g(t) \in T$), то

$$Y_t = f(t)X_{g(t)} + h(t)$$

тоже гауссовский процесс. Более общо, любая конечная детерминированная линейная комбинация $X_{g_j(t)}$ гауссова.

Доказательство: для любых $t_1, \dots, t_n$ все исходные величины, встречающиеся в формулах Y_{t_i} , составляют конечномерный гауссовский вектор V . Тогда $(Y_{t_1}, \dots, Y_{t_n})^\top = AV + b$ с детерминированными A, b .

Для $Y_t = f(t)X_{g(t)} + h(t)$:

$$m_Y(t) = f(t)m_X(g(t)) + h(t),$$

$$R_Y(s, t) = f(s)f(t)R_X(g(s), g(t)).$$

Для $Y_t = \sum_{j=1}^k a_j(t)X_{g_j(t)}$:

$$R_Y(s, t) = \sum_{i,j} a_i(s)a_j(t)R_X(g_i(s), g_j(t)).$$

Нелинейные функции обычно разрушают нормальность: при $X_t \sim N(0, \sigma_t^2)$ и $\sigma_t > 0$ величина $X_t^2/\sigma_t^2 \sim \chi_1^2$, а не нормальна.

6. Задача: какие преобразования гауссовы?

Пусть $(X_t, t \geq 0)$ — центрированный гауссовский процесс с ковариацией $r(s, t)$, причём $r(t, t) > 0$ для всех $t > 0$. Определим:

1. $Y_t = X_t^2$;
2. $Y_t = tX_t$;
3. $Y_t = X_{|\sin t|} - tX_t$;
4. $Y_t = X_{t^5+t} + X_{t^3}$.

Найти все гауссовские процессы.

Решение

1. Нет. При любом $t > 0$, $X_t \sim N(0, r(t, t))$ невырождена, а X_t^2 почти наверное неотрицательна и при этом не является константой. Невырожденная нормальная случайная величина принимает отрицательные значения с положительной вероятностью, следовательно X_t^2 не нормальна.

2. Да. Для любых $t_1, \dots, t_n$ вектор $(t_1 X_{t_1}, \dots, t_n X_{t_n})$ есть диагональное линейное преобразование гауссовского вектора. Ковариация: $R_Y(s, t) = st r(s, t)$.

3. Да. Времена $|\sin t|$ и t неслучайны и неотрицательны. Каждый конечный вектор Y — линейное преобразование конечного гауссовского вектора из значений X в этих временах. Ковариация:

$$R_Y(s, t) = r(|\sin s|, |\sin t|) - tr(|\sin s|, t) - sr(s, |\sin t|) + str(s, t).$$

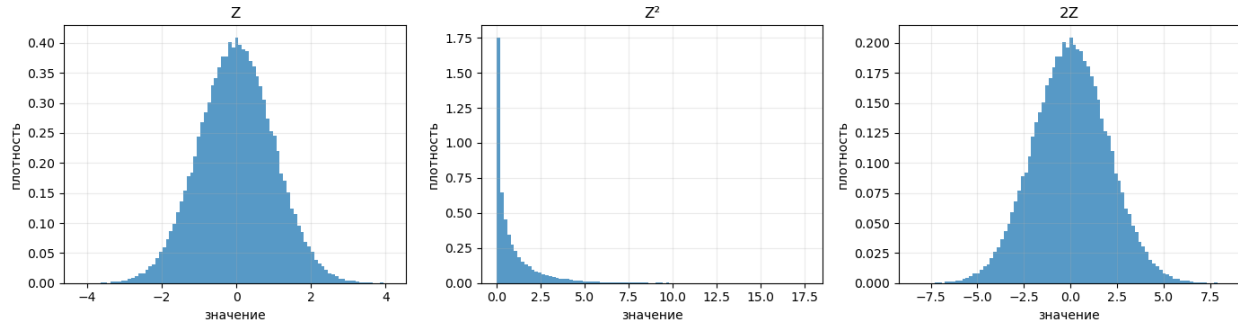
4. Да. Времена $t^5 + t$ и t^3 детерминированны и принадлежат $[0, \infty)$. Конечные линейные комбинации значений X гауссовы. Ковариация:

$$R_Y(s, t) = r(s^5 + s, t^5 + t) + r(s^5 + s, t^3) + r(s^3, t^5 + t) + r(s^3, t^3).$$

Ответ: 2, 3, 4.

6.1 Симуляция 2. Почему возведение в квадрат ломает нормальность

Для фиксированного $t > 0$ нормируем X_t так, что $X_t/\sqrt{r(t, t)} \sim N(0, 1)$. Сравним распределения Z , Z^2 и tZ .



7. Задача: какие функции могут быть ковариациями?

Существует ли центрированный гауссовский процесс с указанной ковариацией $R(s, t)$ на $[0, \infty)$?

1. $R(s, t) = |s - t|$;
2. $R(s, t) = \cos(s - t)$;
3. $R(s, t) = -st$;
4. $R(s, t) = 2$.

Решение

1. Нет. Для $s \neq t$ имеем $R(s, s) = R(t, t) = 0$, но $R(s, t) = |s - t| > 0$. Это противоречит $|R(s, t)|^2 \leq R(s, s)R(t, t) = 0$. Эквивалентно, матрица $\begin{pmatrix} 0 & |s - t| \\ |s - t| & 0 \end{pmatrix}$ имеет отрицательное собственное значение.

2. Да. Возьмём независимые $Z_1, Z_2 \sim N(0, 1)$ и определим $X_t = Z_1 \cos t + Z_2 \sin t$. Тогда $\mathbb{E}X_t = 0$ и

$$\text{cov}(X_s, X_t) = \cos s \cos t + \sin s \sin t = \cos(s - t).$$

Вектор любых значений процесса — линейное преобразование (Z_1, Z_2) , следовательно процесс гауссов.

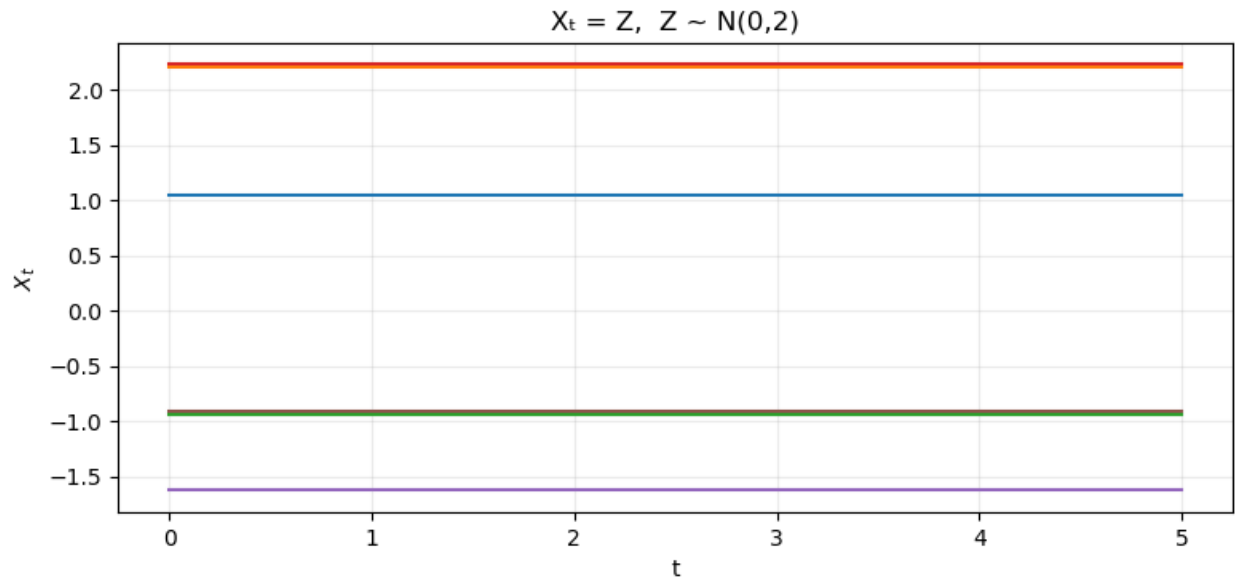
3. Нет. Для $t > 0$, $R(t, t) = -t^2 < 0$, что невозможно для дисперсии.

4. Да. Возьмём $Z \sim N(0, 2)$ и положим $X_t = Z$ для всех t . Тогда $\text{cov}(X_s, X_t) = 2$; все конечномерные распределения совместно нормальны, хотя вырождены.

Ответ: 2, 4.

7.1 Симуляция 4. Постоянная ковариация $R(s, t) = 2$

Если $R(s, t) = 2$, то $\text{Var}(X_t - X_s) = 2 + 2 - 2 \cdot 2 = 0$, то есть $X_s = X_t$ почти наверное для **каждой фиксированной пары** (s, t) . На конечной сетке траектории горизонтальны. Утверждение об одновременном равенстве для всех несчётных t требует отдельно выбирать подходящую версию процесса; построение $X_t = Z$ такую версию даёт.



8. Дополнительные примеры и связь с приложениями

8.1 Броуновское движение

Стандартное броуновское движение — центрированный гауссовский процесс с $R(s, t) = \min(s, t)$, $s, t \geq 0$, и непрерывными траекториями. Его приращения на непересекающихся интервалах независимы, поскольку совместно гауссовы и некоррелированы.

8.2 Стационарные гауссовские процессы

Для стационарного по распределению центрированного гауссовского процесса ковариация зависит только от разности времён: $R(s, t) = k(s - t)$. Примеры: $e^{-|s-t|/\ell}$ и $\cos(\omega(s - t))$. **Не всякая** функция разности допустима: нужна положительная определённость.

8.3 Прогнозирование: условное распределение гауссовского вектора

Если

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left(\begin{pmatrix} m_X \\ m_Y \end{pmatrix}, \begin{pmatrix} K_{XX} & K_{XY} \\ K_{YX} & K_{YY} \end{pmatrix} \right),$$

то при обратимой K_{XX}

$$Y \mid X = x \sim N(m_Y + K_{YX}K_{XX}^{-1}(x - m_X), K_{YY} - K_{YX}K_{XX}^{-1}K_{XY}).$$

Именно эта формула лежит в основе **Gaussian Process Regression** в машинном обучении. При шумных наблюдениях вместо K_{XX} используют $K_{XX} + \sigma_\epsilon^2 I$.

i Приложения

- **ML:** байесовская регрессия с гауссовскими процессами, интервалы неопределённости, байесовская оптимизация.
- **Финансы:** броуновское движение как модель шума; гауссовские модели процентных ставок и факторов риска.
- **Физика:** случайные поля, временные корреляции, тепловые флуктуации.
- **Биология:** гладкие латентные траектории в продольных измерениях и временных рядах.

8.4 Симуляция 5. GP-регрессия: прогноз и неопределённость

Используем ядро $k(s, t) = \exp[-(s - t)^2 / (2\ell^2)]$ и шум наблюдения с дисперсией σ_ε^2 .

