Хитрый лис Иссерлис

Моменты произведений координат гауссовского вектора

Постановка задачи

Пусть случайный вектор $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ распределён нормально:

$$\boldsymbol{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{C}),$$

где $\boldsymbol{\mu} = (\mu_1, ..., \mu_n)^\top$ — вектор средних, $\boldsymbol{C} = \left[C_{ij}\right]$ — ковариационная матрица.

Базовое соотношение:

$$E\left[X_i X_j\right] = \mu_i \mu_j + C_{ij}$$

Задача: найти $E[X_1 X_2 \cdots X_n]$ для произвольного n.

Формула Иссерлиса (предварительная формулировка)

Утверждение: Для гауссовского вектора

$$E[X_1 X_2 \cdots X_n] = h(\boldsymbol{\mu}, \boldsymbol{C})$$

где h — полином, состоящий из слагаемых вида:

$$\mu_{i_1} \mu_{i_2} \cdots \mu_{i_k} \cdot C_{j_1 j_2} C_{j_3 j_4} \cdots C_{j_{n-k-1} j_{n-k}}$$

причём все коэффициенты равны 1.

Доказательство по шагам

Шаг 1: Общая структура — функция от $(\boldsymbol{\mu}, \boldsymbol{C})$

Утверждение: $E[X_1 X_2 \cdots X_n] = h(\boldsymbol{\mu}, \boldsymbol{C})$

Обоснование:

Распределение гауссовского вектора полностью определяется парой $(\boldsymbol{\mu}, \boldsymbol{C})$.

Следовательно, любой момент — функция только этих параметров:

$$E[X_1 X_2 \cdots X_n] = h(\boldsymbol{\mu}, \boldsymbol{C})$$

Размерностный анализ:

- $[E[X_1 \cdots X_n]] = [X]^n$
- $[\mu_i] = [X]$
- $\left[C_{ij}\right] = [X]^2$

$\Rightarrow$ Каждое слагаемое содержит k множителей μ и $\frac{n-k}{2}$ множителей C, где $k + 2 \cdot \frac{n-k}{2} = n$, т.е. k и n одной чётности.

Шаг 2: Структура зависимости от $\boldsymbol{\mu}$

Утверждение: $h(\boldsymbol{\mu}, \boldsymbol{C})$ — полином по $\mu_1, ..., \mu_n$.

Доказательство:

Запишем $X_i = \mu_i + Y_i$, где Y_i — центрированные ($E[Y_i] = 0$).

$$E[X_1 \cdots X_n] = E[(\mu_1 + Y_1)(\mu_2 + Y_2) \cdots (\mu_n + Y_n)]$$

Раскрываем скобки:

$$= \sum_{S \subseteq \{1,\dots,n\}} \prod_{i \in S} \mu_i \cdot E\left[\prod_{j \notin S} Y_j\right]$$

Пример для $n = 3$:

$$\begin{aligned} E[X_1 X_2 X_3] &= \mu_1 \mu_2 \mu_3 \\ &\quad + \mu_1 E[Y_2 Y_3] + \mu_2 E[Y_1 Y_3] + \mu_3 E[Y_1 Y_2] \\ &\quad + E[Y_1 Y_2 Y_3] \end{aligned}$$

Поскольку $E[Y_1 Y_2 Y_3] = 0$ (нечётный момент центрированных), получаем:

$$E[X_1 X_2 X_3] = \mu_1 \mu_2 \mu_3 + \mu_1 C_{23} + \mu_2 C_{13} + \mu_3 C_{12}$$

Шаг 3: От C ничего, кроме C, не зависит — добавим шум

Идея: изменим ковариационную матрицу, добавив независимый шум.

Пусть $Z \sim \mathcal{N}(0,1)$ — независимый от $\boldsymbol{X}$ шум.

Определим новый вектор:

$$\tilde{X}_1 = X_1 + \sqrt{t} \cdot Z, \quad \tilde{X}_i = X_i \text{ для } i \geq 2$$

Как меняется ковариационная матрица?

$$\boldsymbol{C} = \begin{pmatrix} C_{11} & C_{12} & \cdots \\ C_{21} & C_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \to \quad \tilde{\boldsymbol{C}} = \begin{pmatrix} C_{11} + t & C_{12} & \cdots \\ C_{21} & C_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Меняется **только** $C_{11} \to C_{11} + t$.

Ключевое наблюдение:

Момент $M(t) = E\left[\tilde{X}_1^2 \tilde{X}_2 \cdots \tilde{X}_n\right]$ — аналитическая функция по t.

Производная:

$$\frac{dM}{dt}\Big|_{t=0} = \frac{\partial h}{\partial C_{11}}$$

Это показывает, что h **гладко** зависит от элементов $\boldsymbol{C}$, причём зависимость полиномиальная (по размерности).

Шаг 4: Как h зависит от C_{12}? Добавим одинаковый шум к X_1 и X_2

Идея: чтобы изменить **ковариацию** C_{12}, добавим **общий** шум.

Пусть $Z \sim \mathcal{N}(0,1)$ независим от $\boldsymbol{X}$.

Определим:

$$\tilde{X}_1 = X_1 + \sqrt{t} \cdot Z, \quad \tilde{X}_2 = X_2 + \sqrt{t} \cdot Z, \quad \tilde{X}_i = X_i \text{ для } i \geq 3$$

Как меняется ковариационная матрица?

$$\boldsymbol{C} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & \cdots \\ C_{21} & C_{22} & C_{23} & \cdots \\ C_{31} & C_{32} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

↓

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} C_{11} + t & C_{12} + t & C_{13} & \cdots \\ C_{21} + t & C_{22} + t & C_{23} & \cdots \\ C_{31} & C_{32} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Изменились: $C_{11}, C_{22}, C_{12}, C_{21}$ — все на $+t$.

Вычислим производную:

$$\frac{dM}{dt}\ |_{t=0} = \frac{\partial h}{\partial C_{11}} + \frac{\partial h}{\partial C_{22}} + 2\frac{\partial h}{\partial C_{12}}$$

Альтернативно, напрямую дифференцируя:

$$\frac{d}{dt}E\Big[\Big(X_1 + \sqrt{t}Z\Big)\Big(X_2 + \sqrt{t}Z\Big)X_3 \cdots X_n\Big]\ |_{t=0}$$

$$= E\left[\frac{1}{2\sqrt{t}}Z \cdot \Big(X_2 + \sqrt{t}Z\Big)X_3 \cdots + \Big(X_1 + \sqrt{t}Z\Big) \cdot \frac{1}{2\sqrt{t}}Z \cdot X_3 \cdots\right]\ |_{t=0}$$

После вычисления пределов получаем выражение через моменты меньшего порядка.

Сравнивая две формулы, получаем рекуррентное соотношение на коэффициенты h.

Шаг 5: Доказательство того, что все коэффициенты равны 1

Метод: индукция по n с использованием зашумления.

База: $n = 2$

$$E[X_1 X_2] = \mu_1 \mu_2 + C_{12}$$

Коэффициенты: 1 при $\mu_1 \mu_2$ и 1 при C_{12}. ✓

Индуктивный переход:

Рассмотрим момент $M = E[X_1^2 X_2 \cdots X_n]$ (для простоты).

Добавим шум к X_1: $\tilde{X}_1 = X_1 + \sqrt{t}Z$.

$$M(t) = E\Big[\Big(X_1 + \sqrt{t}Z\Big)^2 X_2 \cdots X_n\Big]$$

$$= E[X_1^2 X_2 \cdots X_n] + t \cdot E[Z^2] \cdot E[X_2 \cdots X_n]$$

(используем независимость Z от $X_2, ..., X_n$)

$$= M(0) + t \cdot E[X_2 \cdots X_n]$$

С другой стороны, по формуле $M = h(\boldsymbol{\mu}, \boldsymbol{C})$:

$$M(t) = h\Big(\boldsymbol{\mu}, \boldsymbol{C}|_{C_{11} \to C_{11}+t}\Big)$$

$$\frac{dM}{dt} \ |_{t=0} = \frac{\partial h}{\partial C_{11}}$$

Сравниваем:

$$\frac{\partial h}{\partial C_{11}} = E[X_2 \cdots X_n]$$

По предположению индукции, $E[X_2 \cdots X_n]$ имеет все коэффициенты равные 1.

$\Rightarrow$ В h коэффициент при каждом слагаемом, содержащем C_{11}, равен 1.

Повторяя для всех C_{ij}, получаем: **все** коэффициенты равны 1.

При чём здесь нормальное распределение?

Ключевой момент: Мы использовали нормальность **только** в одном месте —

устойчивость относительно суммирования с шумом:

$$X_i + \sqrt{t} Z \sim \mathcal{N}(\mu_i, C_{ii} + t)$$

Гауссовский вектор **остаётся гауссовским** при добавлении независимого гауссовского шума.

Геометрическая интерпретация:

В пространстве (X_1, Z) преобразование

$$\tilde{X}_1 = X_1 + \sqrt{t} Z = \left(1, \sqrt{t}\right) \cdot \begin{pmatrix} X_1 \\ Z \end{pmatrix}$$

— это **проекция** на направление $\left(1, \sqrt{t}\right)$.

По аксиоме Гермеля–Максвелла:
- Совместное распределение (X_1, Z) изотропно (инвариантно к поворотам)
- Проекция на любое направление — снова гауссова

Для негауссовских распределений:

- Добавление шума **не сохраняет** форму распределения
- Появляются ненулевые кумулянты $\kappa_3, \kappa_4, ...$
- Формула Иссерлиса **неверна**

Теорема Иссерлиса:

Для гауссовского вектора $(X_1, ..., X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$:

$$E[X_1 X_2 \cdots X_n] = \sum_{\text{разбиения}} \prod_{\text{одиночки } \{i\}} \mu_i \cdot \prod_{\text{пары } (i,j)} C_{ij}$$

где сумма берётся по всем способам разбить $\{1, ..., n\}$ на:

- «одиночки» $\to$ вносят μ_i
- «пары» $\to$ вносят C_{ij}

Все коэффициенты равны 1.

Пример: $E[X_1 X_2 X_3 X_4]$ для центрированного случая ($\boldsymbol{\mu} = \mathbf{0}$):

$$E[X_1 X_2 X_3 X_4] = C_{12} C_{34} + C_{13} C_{24} + C_{14} C_{23}$$

Три слагаемых = три способа разбить $\{1, 2, 3, 4\}$ на две пары.