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, $\boldsymbol{C} = [C_{ij}]$ — ковариационная матрица.

Базовое соотношение:

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

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:

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

$\Rightarrow$ 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 $n = 3$:

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

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: Зависимость от диагональных элементов C_{ii} — добавляем шум

Идея: добавим независимый шум $\tilde{Z} \sim \mathcal{N}(0, \sigma^2)$ к одной координате.

Пусть $\tilde{Z}$ независим от $\boldsymbol{X}$. Определим:

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

Вычислим новые ковариации:

$$\mathrm{Var}(\tilde{X}_1) = \mathrm{Var}(X_1) + \mathrm{Var}(\tilde{Z}) = C_{11} + \sigma^2$$

$$\mathrm{Cov}(\tilde{X}_1, X_2) = \mathrm{Cov}(X_1 + \tilde{Z}, X_2) = \mathrm{Cov}(X_1, X_2) + \underbrace{\mathrm{Cov}(\tilde{Z}, X_2)}_{=0} = C_{12}$$

Новая ковариационная матрица:

$$\boldsymbol{C} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & \cdots \\ C_{12} & C_{22} & C_{23} & \cdots \\ C_{13} & C_{23} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \to \quad \tilde{\boldsymbol{C}} = \begin{pmatrix} C_{11} + \sigma^2 & C_{12} & C_{13} & \cdots \\ C_{12} & C_{22} & C_{23} & \cdots \\ C_{13} & C_{23} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Ключевое соотношение:

$$E\left[\tilde{X}_1 X_2 X_3 \cdots X_n\right] = h\left(\boldsymbol{\mu}, \tilde{\boldsymbol{C}}\right)$$

С другой стороны, раскроем напрямую:

$$\begin{aligned} E\left[\tilde{X}_1 X_2 \cdots X_n\right] &= E\left[\left(X_1 + \tilde{Z}\right) X_2 \cdots X_n\right] \\ &= E[X_1 X_2 \cdots X_n] + E\left[\tilde{Z}\right] \cdot E[X_2 \cdots X_n] \\ &= E[X_1 X_2 \cdots X_n] + 0 \end{aligned}$$

Вывод: При изменении $C_{11} \to C_{11} + \sigma^2$ момент **не меняется**!

$$h\left(\boldsymbol{\mu}, \tilde{\boldsymbol{C}}\right) = h(\boldsymbol{\mu}, \boldsymbol{C})$$

$\Rightarrow$ Функция h **не зависит** от C_{11} (и аналогично от любого C_{ii}).

Вывод Шага 3:

В формуле Иссерлиса для $E[X_1 X_2 \cdots X_n]$ с **различными** индексами диагональные элементы C_{ii} **не появляются**.

Они появляются только когда индекс повторяется (например, $E[X_1^2 X_2]$).

Шаг 4: Зависимость от C_{12} — добавляем общий шум к X_1 и X_2

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

Пусть $\tilde{Z} \sim \mathcal{N}(0, \sigma^2)$ независим от $\boldsymbol{X}$.

Определим:

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

Вычислим новые ковариации:

$$\operatorname{Var}\left(\tilde{X}_1\right) = \operatorname{Var}(X_1) + \operatorname{Var}\left(\tilde{Z}\right) = C_{11} + \sigma^2$$

$$\operatorname{Var}\left(\tilde{X}_2\right) = \operatorname{Var}(X_2) + \operatorname{Var}\left(\tilde{Z}\right) = C_{22} + \sigma^2$$

$$\begin{aligned}
\operatorname{Cov}\left(\tilde{X}_1, \tilde{X}_2\right) &= \operatorname{Cov}\left(X_1 + \tilde{Z}, X_2 + \tilde{Z}\right) \\
&= \operatorname{Cov}(X_1, X_2) + \operatorname{Cov}\left(X_1, \tilde{Z}\right) + \operatorname{Cov}\left(\tilde{Z}, X_2\right) + \operatorname{Cov}\left(\tilde{Z}, \tilde{Z}\right) \\
&= C_{12} + 0 + 0 + \sigma^2 = C_{12} + \sigma^2
\end{aligned}$$

$$\operatorname{Cov}\left(\tilde{X}_1, X_3\right) = \operatorname{Cov}\left(X_1 + \tilde{Z}, X_3\right) = C_{13} + 0 = C_{13}$$

Новая ковариационная матрица:

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} C_{11} + \sigma^2 & C_{12} + \sigma^2 & C_{13} & \cdots \\ C_{12} + \sigma^2 & C_{22} + \sigma^2 & C_{23} & \cdots \\ C_{13} & C_{23} & C_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Теперь вычислим момент напрямую:

$$\begin{aligned}
E\left[\tilde{X}_1 \tilde{X}_2 X_3 \cdots X_n\right] &= E\left[\left(X_1 + \tilde{Z}\right)\left(X_2 + \tilde{Z}\right) X_3 \cdots X_n\right] \\
&= E[X_1 X_2 X_3 \cdots X_n] + E\left[X_1 \tilde{Z} X_3 \cdots X_n\right] \\
&\quad + E\left[\tilde{Z} X_2 X_3 \cdots X_n\right] + E\left[\tilde{Z}^2 X_3 \cdots X_n\right]
\end{aligned}$$

Используем независимость $\tilde{Z}$ от $\boldsymbol{X}$:

$$
\begin{aligned}
&= E[X_1 X_2 \cdots X_n] + E\big[\tilde{Z}\big] \cdot E[X_1 X_3 \cdots X_n] \\
&\quad + E\big[\tilde{Z}\big] \cdot E[X_2 X_3 \cdots X_n] + E\big[\tilde{Z}^2\big] \cdot E[X_3 \cdots X_n] \\
&= E[X_1 X_2 \cdots X_n] + 0 + 0 + \sigma^2 \cdot E[X_3 \cdots X_n]
\end{aligned}
$$

Итого:

$$E\big[\tilde{X}_1 \tilde{X}_2 X_3 \cdots X_n\big] = E[X_1 X_2 \cdots X_n] + \sigma^2 \cdot E[X_3 \cdots X_n]$$

С другой стороны, по функции h:

$$h\Big(\boldsymbol{\mu}, \tilde{\boldsymbol{C}}\Big) = h(\boldsymbol{\mu}, \boldsymbol{C}) + \frac{\partial h}{\partial C_{11}}\sigma^2 + \frac{\partial h}{\partial C_{22}}\sigma^2 + \frac{\partial h}{\partial C_{12}}\sigma^2 + \frac{\partial h}{\partial C_{21}}\sigma^2 + O(\sigma^4)$$

Но из Шага 3: $\frac{\partial h}{\partial C_{11}} = \frac{\partial h}{\partial C_{22}} = 0$.

И $C_{12} = C_{21}$, поэтому:

$$h\Big(\boldsymbol{\mu}, \tilde{\boldsymbol{C}}\Big) = h(\boldsymbol{\mu}, \boldsymbol{C}) + 2\frac{\partial h}{\partial C_{12}}\sigma^2 + O(\sigma^4)$$

Сравнивая:

$$2\frac{\partial h}{\partial C_{12}}\sigma^2 = \sigma^2 \cdot E[X_3 \cdots X_n]$$

$$\Rightarrow \quad \frac{\partial h}{\partial C_{12}} = \frac{1}{2}E[X_3 \cdots X_n]$$

Вывод Шага 4:

$$\frac{\partial h}{\partial C_{12}} = \frac{1}{2}E[X_3 X_4 \cdots X_n]$$

Это означает: функция h **линейна** по каждому C_{ij}, и коэффициент при C_{12} пропорционален моменту **оставшихся** переменных.

Шаг 5.1: Все коэффициенты равны 1 — метод специальных подстановок

Идея: выберем **конкретные** значения $\boldsymbol{\mu}$ и $\boldsymbol{C}$, чтобы «выключить» все слагаемые кроме одного и проверить его коэффициент.

Подстановка 1: Проверка коэффициента при $C_{12}C_{34}$

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

По формуле (которую проверяем):

$$E[X_1 X_2 X_3 X_4] = \alpha \cdot C_{12}C_{34} + \beta \cdot C_{13}C_{24} + \gamma \cdot C_{14}C_{23}$$

Специальная ковариационная матрица:

Положим:

$$\boldsymbol{C} = \begin{pmatrix} 1 & c & 0 & 0 \\ c & 1 & 0 & 0 \\ 0 & 0 & 1 & c \\ 0 & 0 & c & 1 \end{pmatrix}$$

где $c \in (0, 1)$ — параметр.

При такой матрице:
- $C_{12} = C_{34} = c$
- $C_{13} = C_{14} = C_{23} = C_{24} = 0$

Тогда:

$$E[X_1 X_2 X_3 X_4] = \alpha \cdot c^2 + \beta \cdot 0 + \gamma \cdot 0 = \alpha c^2$$

Вычислим напрямую:

Вектор (X_1, X_2) независим от (X_3, X_4).

$$E[X_1 X_2 X_3 X_4] = E[X_1 X_2] \cdot E[X_3 X_4] = C_{12} \cdot C_{34} = c^2$$

Сравнивая: $\alpha c^2 = c^2 \Rightarrow \alpha = 1 \checkmark$

Подстановка 2: Проверка коэффициента при $\mu_1 \mu_2 C_{34}$
Рассмотрим $E[X_1 X_2 X_3 X_4]$ для нецентрированного случая.

Специальные параметры:

$$\boldsymbol{\mu} = (a, a, 0, 0)^\top, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & c \\ 0 & 0 & c & 1 \end{pmatrix}$$

При этом:
- $\mu_1 = \mu_2 = a$, $\mu_3 = \mu_4 = 0$
- $C_{12} = C_{13} = C_{14} = C_{23} = C_{24} = 0$
- $C_{34} = c$

Ожидаемая форма:

$$E[X_1 X_2 X_3 X_4] = \beta \cdot \mu_1 \mu_2 C_{34} + ... = \beta a^2 c + ...$$

Вычислим напрямую:

(X_1, X_2) независимы от (X_3, X_4), и X_1, X_2 независимы между собой.

$$\begin{aligned} E[X_1 X_2 X_3 X_4] &= E[X_1] \cdot E[X_2] \cdot E[X_3 X_4] \\ &= a \cdot a \cdot C_{34} = a^2 c \end{aligned}$$

Сравнивая: $\beta a^2 c = a^2 c \Rightarrow \beta = 1 \checkmark$

Общая схема подстановок

Алгоритм проверки коэффициента при слагаемом $\mu_{i_1} \cdots \mu_{i_k} C_{j_1 j_2} \cdots C_{j_{m-1} j_m}$:

1. Положить $\mu_i = a$ для $i \in \{i_1, ..., i_k\}$, иначе $\mu_i = 0$
2. Положить $C_{j_{2\ell-1}, j_{2\ell}} = c$ для нужных пар, остальные $C_{ij} = 0$ для $i \neq j$
3. Вычислить $E[X_1 \cdots X_n]$ напрямую (используя независимость)
4. Сравнить с коэфф $\cdot\, a^k c^{\frac{n-k}{2}}$

$\Rightarrow$ Коэффициент всегда равен 1

Шаг 5.2: Метод индукция по n с использованием зашумления

Если шум $Z \sim N(0, 1)$

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: $\tilde{X}_1 = X_1 + \sqrt{t} Z$

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

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

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

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

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

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

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

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 в двух местах:

1. **Шаг 1:** распределение полностью определяется $(\boldsymbol{\mu}, \boldsymbol{C})$
2. **Шаги 3–4:** устойчивость относительно добавления гауссовского шума:

$$X_i + \tilde{Z} \sim \mathcal{N}(\mu_i, C_{ii} + \sigma^2)$$

Гауссовский вектор **остаётся гауссовским** при добавлении независимого гауссовского шума — это **поворот** в расширенном пространстве.

Связь с аксиомой Гермеля–Максвелла:

В пространстве $\left(X_1, \tilde{Z}\right)$ преобразование

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

— это **проекция** на направление $(1,1)$.

По аксиоме Гермеля–Максвелла:
- Совместное распределение $\left(X_1, \tilde{Z}\right)$ изотропно
- Проекция на любое направление — снова гауссова

Для негауссовских распределений:

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

Теорема Иссерлиса формально:

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

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

где сумма берётся по всем способам разбить $\{1, \ldots, n\}$ на:
- «одиночки» $\to$ вносят μ_i
- «пары» $\to$ вносят C_{ij}

Все коэффициенты равны 1