

Двумерное гауссово распределение

Содержание

7.1.	Физическая задача: два показания калориметра	91
7.2.	От двух одномерных гауссов к двумерному	91
7.3.	Многомерный гаусс и расстояние Махаланобиса	93
7.4.	Сколько вероятности находится внутри эллипса	94
7.5.	Объединение измерений	96
7.6.	Инвариантная масса двух частиц	97
7.7.	Где заканчивается гауссова геометрия	98
7.8.	Задачи	99
	Литература	100

Аннотация

Совместное распределение двух измеряемых величин содержит больше информации, чем две ошибки, выписанные по отдельности. Ковариационная матрица задаёт масштаб и направление флуктуаций, а её обратная матрица позволяет измерить расстояние от точки до среднего. В этой главе мы построим двумерный гаусс, разберём вероятностные эллипсы и применим полную ковариацию к объединению измерений и восстановлению инвариантной массы.

7.1. Физическая задача: два показания калориметра

Рассмотрим ансамбль измерений событий с одной и той же кинематикой. Истинные энергии двух фотонов равны 60 и 40 ГэВ. Будем считать, что шкала энергии откалибрована без смещения, а отклик калориметра приближённо гауссов. Тогда средний вектор измеренных энергий равен

$$\mu = \begin{pmatrix} 60 \\ 40 \end{pmatrix} \text{ ГэВ.} \quad (7.1)$$

Дисперсии измеренных энергий и их ковариация собраны в матрицу

$$V = \begin{pmatrix} 1.44 & 0.36 \\ 0.36 & 1.00 \end{pmatrix} \text{ ГэВ}^2. \quad (7.2)$$

Вне диагонали стоит положительное число: часть ошибки у двух энергий общая. Например, обе зависят от одной калибровки шкалы энергии.

Пусть в одном событии получено

$$\mathbf{x} = \begin{pmatrix} 61.2 \\ 41.0 \end{pmatrix} \text{ ГэВ.} \quad (7.3)$$

Каждая энергия отличается от среднего ровно на своё стандартное отклонение. Этого ещё недостаточно, чтобы оценить совместное отклонение. Для него нужна квадратичная форма

$$D^2 = (\mathbf{x} - \mu)^T V^{-1} (\mathbf{x} - \mu) = 1.54. \quad (7.4)$$

Число D^2 безразмерно. Оно учитывает и масштабы двух ошибок, и направление, в котором они коррелируют. Контур, проходящий через нашу точку, охватывает около 54% двумерного гауссова распределения. Для области с вероятностью 68.3% потребуется $D^2 \leq 2.30$.

Откуда взялись эти числа и почему привычное условие $D^2 \leq 1$ здесь даёт не 68.3%, а только 39.3%? Для ответа надо построить двумерный гаусс и посмотреть на геометрию его ковариационной матрицы.

7.2. От двух одномерных гауссов к двумерному**7.2.1. Независимые величины**

Пусть

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2), \quad Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \quad (7.5)$$

и X и Y независимы. Совместная плотность равна произведению двух одномерных плотностей:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left[-\frac{1}{2}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right]. \quad (7.6)$$

Введём безразмерные отклонения

$$u = \frac{x - \mu_X}{\sigma_X}, \quad v = \frac{y - \mu_Y}{\sigma_Y}. \quad (7.7)$$

Плотность постоянна при постоянном значении

$$u^2 + v^2 = c. \quad (7.8)$$

В координатах (u, v) это окружность. В исходных координатах получается эллипс с полуосями $\sqrt{c}\sigma_X$ и $\sqrt{c}\sigma_Y$. Его оси совпадают с координатными: отклонение X ничего не сообщает об отклонении Y .

7.2.2. Корреляция поворачивает эллипс

Ковариационная матрица двух величин имеет вид

$$V = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}, \quad |\rho| < 1. \quad (7.9)$$

Смешанный член в совместной плотности появляется вместе с корреляцией:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{u^2 - 2\rho uv + v^2}{1-\rho^2}\right]. \quad (7.10)$$

При $\rho = 0$ уравнение 7.10 переходит в произведение двух одномерных гауссов. При положительной корреляции большие значения X чаще сопровождаются большими значениями Y , поэтому эллипс наклонён вверх. Отрицательная корреляция меняет направление наклона.

Предел $|\rho| = 1$ в формулу плотности не входит. В этом случае определитель ковариационной матрицы обращается в нуль, а случайные точки лежат на прямой. Двумерной плотности относительно площади уже нет.

7.2.3. Обратная ковариационная матрица

Для матрицы 7.9

$$\det V = \sigma_X^2\sigma_Y^2(1-\rho^2), \quad (7.11)$$

а обратная матрица равна

$$V^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1/\sigma_X^2 & -\rho/(\sigma_X\sigma_Y) \\ -\rho/(\sigma_X\sigma_Y) & 1/\sigma_Y^2 \end{pmatrix}. \quad (7.12)$$

Обозначим отклонение от среднего через

$$\delta = \mathbf{x} - \boldsymbol{\mu} = \begin{pmatrix} x - \mu_X \\ y - \mu_Y \end{pmatrix}. \quad (7.13)$$

Тогда показатель экспоненты можно записать одной строкой:

$$\delta^T V^{-1} \delta = \frac{u^2 - 2\rho uv + v^2}{1 - \rho^2}. \quad (7.14)$$

Формула через ρ и матричная формула описывают одну плотность. Матрица V задаёт флуктуации, а V^{-1} оценивает отклонение с учётом этих флуктуаций.

7.2.4. Проверка коэффициента ρ

Построим коррелированные величины из двух независимых стандартных гауссов:

$$Z_1, Z_2 \sim \mathcal{N}(0, 1), \quad \text{cov}(Z_1, Z_2) = 0, \quad (7.15)$$

$$X = \mu_X + \sigma_X Z_1, \quad Y = \mu_Y + \sigma_Y \left(\rho Z_1 + \sqrt{1 - \rho^2} Z_2 \right). \quad (7.16)$$

Такой способ генерации пригодится и в методе Монте-Карло. Пока проверим, что параметры в нём имеют заявленный смысл.

Вывод

7.2.4.1. Средние, дисперсии и ковариация

Из уравнения 7.16 сразу следуют средние

$$\mathbb{E}[X] = \mu_X, \quad \mathbb{E}[Y] = \mu_Y. \quad (7.17)$$

Дисперсия X равна σ_X^2 . Для второй величины получаем

$$\begin{aligned} \text{Var}(Y) &= \sigma_Y^2 [\rho^2 \text{Var}(Z_1) + (1 - \rho^2) \text{Var}(Z_2)] \\ &= \sigma_Y^2 [\rho^2 + (1 - \rho^2)] = \sigma_Y^2. \end{aligned} \quad (7.18)$$

Смешанное среднее даёт

$$\begin{aligned} \text{cov}(X, Y) &= \sigma_X \sigma_Y \text{cov}\left(Z_1, \rho Z_1 + \sqrt{1 - \rho^2} Z_2\right) \\ &= \rho \sigma_X \sigma_Y. \end{aligned} \quad (7.19)$$

Поэтому параметр ρ в уравнении 7.10 равен коэффициенту корреляции:

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (7.20)$$

7.3. Многомерный гаусс и расстояние Махаланобиса

Для случайного вектора размерности d со средним μ и невырожденной ковариационной матрицей V плотность имеет вид [1]

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sqrt{\det V}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu)^T V^{-1}(\mathbf{x} - \mu)\right]. \quad (7.21)$$

Квадратичная форма

$$D^2 = (\mathbf{x} - \mu)^T V^{-1}(\mathbf{x} - \mu) \quad (7.22)$$

называется квадратом расстояния Махаланобиса. Обычное евклидово расстояние считает один ГэВ по каждой оси одинаковым отклонением. Расстояние 7.22 сначала делит отклонения на их характерные масштабы и одновременно учитывает корреляции.

7.3.1. Главные оси

Поскольку ковариационная матрица симметрична, её можно разложить как

$$V = R\Lambda R^T, \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_d). \quad (7.23)$$

Столбцы ортогональной матрицы R являются собственными векторами V . Перейдём к координатам

$$\mathbf{z} = \Lambda^{-1/2} R^T (\mathbf{x} - \mu). \quad (7.24)$$

В них ковариационная матрица единична, а расстояние принимает обычный вид

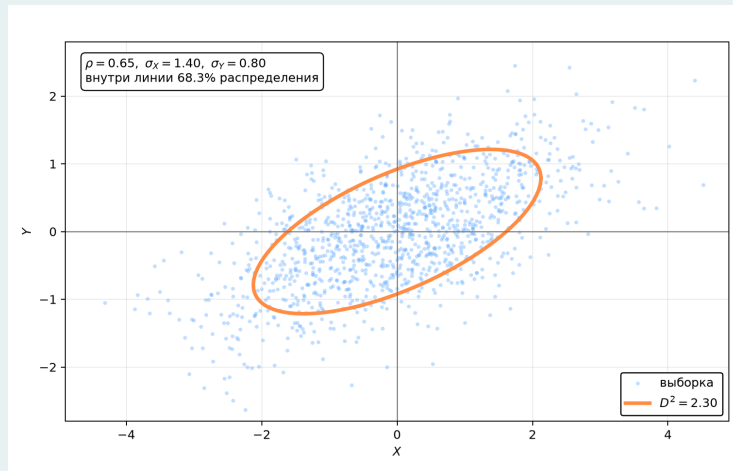
$$D^2 = \mathbf{z}^T \mathbf{z} = \sum_{i=1}^d z_i^2. \quad (7.25)$$

Отсюда сразу читается геометрия контура $D^2 = c$. В двух измерениях это эллипс. Его главные оси направлены вдоль собственных векторов V , а длины полуосей равны

$$a_i = \sqrt{c\lambda_i}. \quad (7.26)$$

Коэффициент корреляции поворачивает собственные векторы относительно исходных осей. Собственные значения определяют ширины распределения в новых направлениях.

7.3.2. Интерактив: корреляция и эллипс



Интерактив. корреляция и эллипс

Меняйте ρ , σ_x и σ_y . Точки показывают выборку из двумерного гаусса, линия — эллипс $D^2 = 2.30$.



При изменении ρ меняются сразу направление и длины осей. Матрица с теми же диагональными элементами, но другой ковариацией задаёт другое совместное распределение.

7.4. Сколько вероятности находится внутри эллипса

В одном измерении фраза «в пределах одной сигмы» понятна:

$$P(|Z| \leq 1) = 0.683, \quad Z \sim \mathcal{N}(0, 1). \quad (7.27)$$

После появления второй оси требуется уточнить область. Для двумерного гаусса стандартизованные координаты независимы, поэтому

$$D^2 = Z_1^2 + Z_2^2 \sim \chi_2^2. \quad (7.28)$$

Плотность χ_2^2 особенно проста:

$$f_{\chi_2^2}(q) = \frac{1}{2}e^{-q/2}, \quad q \geq 0. \quad (7.29)$$

Интегрируя её от нуля до c , получаем вероятность внутри эллипса:

$$P(D^2 \leq c) = 1 - e^{-c/2}. \quad (7.30)$$

Для заданной вероятности P граница находится из обратной формулы

$$c = -2 \ln(1 - P). \quad (7.31)$$

Значение $c = 1$ в двух измерениях даёт

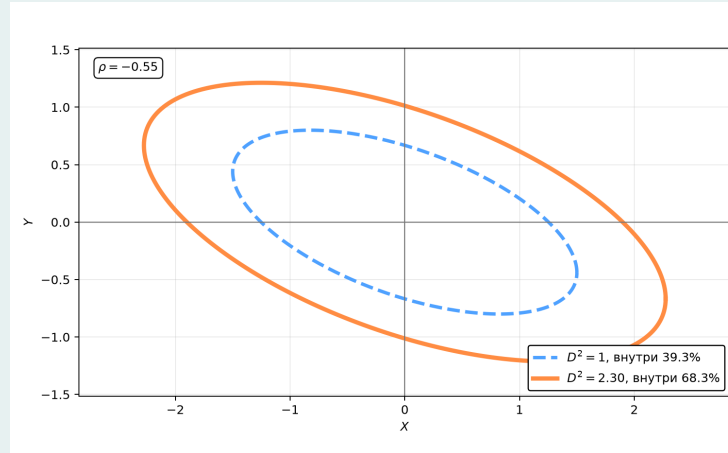
$$P(D^2 \leq 1) = 1 - e^{-1/2} = 0.393. \quad (7.32)$$

В таблице 7.1 приведены уровни, которые часто встречаются на двумерных графиках.

Таблица 7.1. Вероятностные уровни двумерного гауссова распределения

Вероятность внутри контура	Граница c
39.3%	1.00
68.3%	2.30
90%	4.61
95%	5.99
99%	9.21

7.4.1. Интерактив: вероятность и размер контура



Интерактив. вероятность и размер контура

Сплошная линия ограничивает выбранную долю распределения. Пунктир показывает $D^2 = 1$, который в двух измерениях содержит 39.3% вероятности.



7.4.2. Возвращение к калориметру

Для измерения 7.3 мы нашли $D^2 = 1.54$. Уравнение 7.30 даёт

$$P(D^2 \leq 1.54) = 1 - e^{-1.54/2} = 0.537. \quad (7.33)$$

Иными словами, эллипс через эту точку содержит около 54% распределения. Точка лежит внутри области 68.3%, поскольку $1.54 < 2.30$. Положительная корреляция здесь существенна: оба измерения отклонились в одну сторону, то есть вдоль вытянутой оси облака.

7.4.3. Распределение данных и пространство параметров

Эллипсы этой главы описывают случайный вектор $\mathbf{X}$ при известных μ и V . В статьях похожие контуры часто рисуют в пространстве оцениваемых параметров вокруг полученного минимума. Одинаковая квадратичная форма ещё не делает их одним объектом.

Для параметров вероятностная интерпретация зависит от того, как построен интервал: от распределения оценки в частотном подходе или от апостериорной плотности в байесовском. Значения 2.30, 4.61 и 5.99 применимы к двумерной квадратичной гауссовой задаче. Условия, при которых ими можно пользоваться для параметров, мы обсудим вместе с правдоподобием и χ^2 .

7.5. Объединение измерений

Пусть n приборов измеряют одну величину μ . Результаты собраны в вектор $\mathbf{x}$, а их ковариационная матрица равна V . Если совместное распределение гауссово, значение μ , при котором полученные данные имеют наибольшую плотность, минимизирует

$$Q(\mu) = (\mathbf{x} - \mu\mathbf{1})^T V^{-1} (\mathbf{x} - \mu\mathbf{1}). \quad (7.34)$$

Здесь $\mathbf{1}$ — вектор из единиц. Дифференцирование по μ даёт

$$\hat{\mu} = \frac{\mathbf{1}^T V^{-1} \mathbf{x}}{\mathbf{1}^T V^{-1} \mathbf{1}}, \quad \text{Var}(\hat{\mu}) = \frac{1}{\mathbf{1}^T V^{-1} \mathbf{1}}. \quad (7.35)$$

Это несмещённая линейная комбинация с наименьшей дисперсией. При независимых измерениях V диагональна, и формула принимает знакомый вид

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i / \sigma_i^2}{\sum_{i=1}^n 1 / \sigma_i^2}, \quad \text{Var}(\hat{\mu}) = \frac{1}{\sum_{i=1}^n 1 / \sigma_i^2}. \quad (7.36)$$

Для двух измерений с ковариацией c_{12} веса можно выписать явно:

$$\hat{\mu} = w_1 x_1 + w_2 x_2, \quad w_1 = \frac{\sigma_2^2 - c_{12}}{\sigma_1^2 + \sigma_2^2 - 2c_{12}}, \quad w_2 = \frac{\sigma_1^2 - c_{12}}{\sigma_1^2 + \sigma_2^2 - 2c_{12}}. \quad (7.37)$$

Сумма весов равна единице. Один из них может оказаться отрицательным, если корреляция велика, а точности измерений заметно различаются. Отрицательный вес помогает оценить общую флуктуацию по второму измерению и частично вычесть её из более точного результата.

7.6. Инвариантная масса двух частиц

Возьмём ещё один пример из лекции. Для двух безмассовых частиц с энергиями E_1, E_2 и углом θ между импульсами

$$m^2 = 2E_1 E_2 (1 - \cos \theta). \quad (7.38)$$

Вектор измеряемых величин и его ковариационная матрица равны

$$\mathbf{x} = \begin{pmatrix} E_1 \\ E_2 \\ \theta \end{pmatrix}, \quad V_{\mathbf{x}} = \begin{pmatrix} 1.44 & 0.36 & 0 \\ 0.36 & 1.00 & 0 \\ 0 & 0 & 2.5 \cdot 10^{-5} \end{pmatrix}. \quad (7.39)$$

В энергетическом блоке числа указаны в ГэВ², угловая дисперсия — в рад². Возьмём

$$E_1 = 60.0 \text{ ГэВ}, \quad E_2 = 40.0 \text{ ГэВ}, \quad \theta = 0.200 \text{ рад}. \quad (7.40)$$

Градиент квадрата массы равен

$$\nabla m^2 = \begin{pmatrix} 2E_2(1 - \cos \theta) \\ 2E_1(1 - \cos \theta) \\ 2E_1 E_2 \sin \theta \end{pmatrix}. \quad (7.41)$$

Вывод

7.6.0.1. Численный расчёт

Сначала найдём саму массу:

$$m^2 = 2 \cdot 60 \cdot 40 (1 - \cos 0.200) = 95.680 \text{ ГэВ}^2, \quad m = 9.782 \text{ ГэВ}. \quad (7.42)$$

В заданной точке численный градиент равен

$$\nabla m^2 = \begin{pmatrix} 1.595 \\ 2.392 \\ 953.613 \end{pmatrix}. \quad (7.43)$$

По формуле распространения ковариации из главы 6

$$\text{Var}(m^2) \simeq (\nabla m^2)^T V_{\mathbf{x}} (\nabla m^2) = 34.864 \text{ ГэВ}^4. \quad (7.44)$$

Отсюда

$$\sigma_{m^2} = 5.905 \text{ ГэВ}^2, \quad \sigma_m \simeq \frac{\sigma_{m^2}}{2m} = 0.302 \text{ ГэВ}. \quad (7.45)$$

Результат можно записать как

$$m = (9.78 \pm 0.30) \text{ ГэВ}. \quad (7.46)$$

Ковариация двух энергий даёт в дисперсию m^2 вклад

$$2 \frac{\partial m^2}{\partial E_1} \frac{\partial m^2}{\partial E_2} \text{cov}(E_1, E_2) = 2.75 \text{ ГэВ}^4. \quad (7.47)$$

Если этот член отбросить, получится $\sigma_m = 0.290 \text{ ГэВ}$ вместо 0.302 ГэВ . Та же положительная корреляция, которая наклоняла эллипс двух энергий, увеличивает ошибку массы: обе производные по энергиям положительны.

7.7. Где заканчивается гауссова геометрия

Многомерный гаусс полностью определяется средним и ковариационной матрицей. Для произвольного распределения этих двух объектов уже недостаточно. Два распределения могут иметь одинаковые μ и V , но разные хвосты и разную вероятность далёких отклонений.

Линейное распространение ковариации хорошо описывает малые флуктуации около рабочей точки. Сильная нелинейность, асимметричные входные распределения и физические границы требуют распространения полного распределения. Один из способов сделать это — сгенерировать входные величины с заданными связями и пропустить их через функцию результата.

Отдельная численная трудность возникает при $|\rho| \simeq 1$. Матрица V становится почти вырожденной, и результат обращения чувствителен к малым изменениям её элементов. Это означает, что два измерения содержат почти одну и ту же информацию. Печатать больше знаков в V^{-1} здесь бесполезно: следует проверить модель общей ошибки.

Итоги главы

- Ковариационная матрица задаёт масштабы и направления флуктуаций многомерного гаусса.
- Квадрат расстояния Махаланобиса учитывает единицы, дисперсии и корреляции.
- В двух измерениях область $D^2 \leq 1$ содержит 39.3% вероятности; для 68.3% нужна граница $D^2 = 2.30$.
- Объединение измерений и распространение ошибок требуют полной ковариационной матрицы.

7.8. Задачи

Задача 1. Вероятность внутри эллипса

Для двумерного гауссова распределения

$$D^2 = (\mathbf{X} - \boldsymbol{\mu})^T V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_2^2.$$

1. Получите плотность D^2 и покажите, что

$$P(D^2 \leq c) = 1 - e^{-c/2}.$$

2. Найдите c для областей с вероятностями 68.3%, 90% и 95%.
3. Вычислите вероятность внутри контура $D^2 = 1$.
4. Сравните этот результат с одномерной вероятностью $P(|Z| \leq 1)$ и объясните различие.

Задача 2. Геометрия ковариационной матрицы

Даны средний вектор и ковариационная матрица двух измеренных энергий:

$$\boldsymbol{\mu} = \begin{pmatrix} 60 \\ 40 \end{pmatrix} \text{ ГэВ}, \quad V = \begin{pmatrix} 1.44 & 0.36 \\ 0.36 & 1.00 \end{pmatrix} \text{ ГэВ}^2.$$

1. Найдите стандартные отклонения и коэффициент корреляции.
2. Вычислите собственные значения и собственные векторы V .
3. Найдите длины и направления полуосей эллипса $D^2 = 2.30$.
4. Для точки $\mathbf{x} = (61.2, 41.0)^T$ ГэВ вычислите D^2 и определите, лежит ли она внутри области с вероятностью 68.3%.

Задача 3. Объединение двух измерений

Одна величина измерена двумя способами:

$$x_1 = 10.2 \pm 0.8, \quad x_2 = 11.1 \pm 0.4.$$

1. Найдите взвешенное среднее и его стандартное отклонение, считая измерения независимыми.
2. Повторите расчёт для коэффициента корреляции $\rho = 0.6$.
3. Выпишите оба веса во втором случае и объясните знак каждого из них.
4. Проверьте, что ковариационная матрица положительно определена.

Задача 4. Инвариантная масса

Для двух безмассовых частиц

$$m^2 = 2E_1E_2(1 - \cos \theta).$$

Пусть $E_1 = 60$ ГэВ, $E_2 = 40$ ГэВ, $\theta = 0.200$ рад, а ковариационная матрица величин (E_1, E_2, θ) равна

$$V = \begin{pmatrix} 1.44 & 0.36 & 0 \\ 0.36 & 1.00 & 0 \\ 0 & 0 & 2.5 \cdot 10^{-5} \end{pmatrix}.$$

1. Найдите градиент m^2 и вычислите σ_{m^2} .
2. Получите σ_m линейным распространением ошибки.
3. Повторите расчёт при $\text{cov}(E_1, E_2) = 0$ и сравните ответы.
4. Исследуйте предел $\theta \ll 1$: получите приближённые выражения для m и его производных.

Литература

[1] Cowan, G. *Statistical Data Analysis*. Oxford University Press. 1998.