Лекция 16

ГАУССОВСКИЕ СЛУЧАЙНЫЕ ВЕКТОРЫ

1. Двумерные гауссовские случайные векторы

В лекции 15 мы рассматривали двумерную стандартную гауссовскую плотность. Случайный вектор, который имеет такую плотность, мы называли стандартным гауссовским вектором.

Как и в случае случайных величин (см. §8.2, лекция 14), мы определим общие гауссовские векторы, но при этом будем пользоваться другим определением. Позднее мы докажем, что новое определение эквивалентно определению из раздела 15.4.

Определение 1. Случайный вектор $\mathbf{X} = (X_1, X_2)'$ называется *гауссовским* случайным вектором, если линейная комбинация $c_1X_1 + c_2X_2$ является гауссовской случайной величиной при любых c_1 и c_2 , $|c_1| + |c_2| \neq 0$.

Напомним, что таким свойством обладают стандартные гауссовские случайные векторы (см. §4.2, лекция 15).

1.1. Простейшие свойства. Непосредственно из определения 1 вытекают такие свойства гауссовских векторов:

- (i) гауссовские векторы существуют: например, если X_1 и X_2 независимые стандартные гауссовские случайные величины, то вектор $\mathbf{X}=(X_1,X_2)'$ является гауссовским;
- (ii) если **X** гауссовский случайный вектор, то X_1 и X_2 являются гауссовскими случайными величинами; (не обязательно независимыми!)
- (iii) если ${\bf X}$ гауссовский случайный вектор, то X_1+X_2 является гауссовской случайной величиной.

Доказательство. (i) Сначала докажем, что $Y \stackrel{\text{def}}{=} cX \in \mathcal{N}\left(0,c^2\sigma^2\right)$ для любого $c \neq 0$, если $X \in \mathcal{N}\left(0,\sigma^2\right)$. Действительно, из теоремы 15.2 получаем $f_Y(y) = \frac{1}{|c|} f_X(y/c)$ или $f_Y(y) = \frac{1}{\sqrt{2\pi c^2\sigma^2}} e^{-(y/c)^2/2\sigma^2} = \varphi_{0,c^2\sigma^2}(y)$.

Теперь покажем, что $cX_1 + X_2 \in \mathcal{N}(0, 1 + c^2)$ для любого $c \in \mathbf{R}$. Это утверждение тривиально для c = 0. Докажем его и для $c \neq 0$. Так как $cX_1 \in \mathcal{N}(0, c^2)$, то из формулы свертки (15.6) получаем

$$f_{cX_1+X_2}(x) = \frac{1}{2\pi|c|} \int_{-\infty}^{\infty} e^{-v^2/2c^2} \cdot e^{-(x-v)^2/2} dv.$$

Положим $\sigma^2 = \frac{c^2}{1+c^2}$. Тогда

$$\frac{v^2}{c^2} + (x - v)^2 - \frac{x^2}{1 + c^2} = \frac{v^2}{\sigma^2} - 2xv + \sigma^2 x^2 = \left(\frac{v}{\sigma} - \sigma x\right)^2 = \frac{1}{\sigma^2} (v - \sigma^2 x)^2.$$

 $^{^0\}mathrm{Printed}$ from the file [16-Gauss-vectors.tex] on 23.3.2016

Поэтому

$$\begin{split} f_{cX_1+X_2}(x) &= \frac{1}{2\pi|c|} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{v^2}{\sigma^2} + (x-v)^2 \pm \frac{x^2}{1+c^2}\right)} \, dv \\ &= \varphi_{0,1+c^2}(x) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(v-\sigma^2x)^2} \, dv \\ &= \varphi_{0,1+c^2}(x) \int_{-\infty}^{\infty} \varphi_{\sigma^2x,\sigma^2}(v) \, dv = \varphi_{0,1+c^2}(x), \end{split}$$

что и требовалось доказать.

Теперь легко закончить доказательство утверждения (i). Пусть, например, $c_2 \neq 0$. Тогда $c_1X_1+c_2X_2=c_2(cX_1+X_2)$, где $c=c_1/c_2$. Поскольку $cX_1+X_2\in \mathcal{N}\left(0,1+c^2\right)$, то $c_2(cX_1+X_2)\in \mathcal{N}\left(0,c_2^2(1+c^2)\right)=\mathcal{N}\left(0,c_1^2+c_2^2\right)$ по доказанному выше. Значит $c_1X_1+c_2X_2$ является гауссовской случайной величиной с параметрами 0 и $c_1^2+c_2^2$.

- (ii) Достаточно выбрать $c_1=1, c_2=0$ или $c_1=0, c_2=1$ в определении 1.
- (iii) Достаточно выбрать $c_1 = c_2 = 1$ в определении 1. \square
- **1.2.** Вектор негауссовский, а координаты гауссовские. Если координаты вектора являются гауссовскими, то сам вектор не обязательно является гауссовским (свойство (ii) обратить нельзя). Действительно, пусть $X_1 \in \mathcal{N}(0,1)$, а β симметричная случайная величина Бернулли, то есть $\mathsf{P}(\beta=-1) = \mathsf{P}(\beta=+1) = \frac{1}{2}$. Предположим, что β не зависит от X_1 и положим $X_2 = \beta X_1$. Так как

$$\{X_2 < v\} = \{X_2 < v, \beta = -1\} \cup \{X_2 < v, \beta = +1\}$$

= \{-X_1 < v, \beta = -1\} \cup \{X_1 < v, \beta = +1\},

то в силу независимости

$$P(X_2 < v) = \frac{1}{2}P(-X_1 < v) + \frac{1}{2}P(X_1 < v) = PP(X_1 < v)$$

в силу симметричности гауссовского распределения (см. §8.1, лекция 14). С другой стороны,

$$P(X_1 + X_2 = 0) = P(\beta = -1) = \frac{1}{2},$$

откуда мы заключаем, что $X_1 + X_2$ негауссовская случайная величина. $\mathbb O$ В силу свойства (iii) это означает, что вектор $\mathbf X$ не гауссовский.

1.3. Ковариационная матрица. Поскольку координаты двумерного гауссовского вектора являются гауссовскими случайными величинами, то они имеют математические ожидания и дисперсии. Пусть $\boldsymbol{\mu}$ — вектор-столбец математических ожиданий вектора \mathbf{X} : $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, где $\mu_1 = \mathsf{E}[X_1]$ и $\mu_2 = \mathsf{E}[X_2]$. Положим $\sigma_1^2 = \mathsf{var}[X_1]$ и $\sigma_2^2 = \mathsf{var}[X_2]$, а ковариацию между X_1 и X_2 обозначим через $\rho = \mathsf{E}[(X_1 - \mu_1)(X_2 - \mu_2)]$. Матрица

$$\boldsymbol{\Lambda} \stackrel{\mathrm{def}}{=} \begin{pmatrix} \operatorname{cov}\left[X_1, X_1\right] & \operatorname{cov}\left[X_1, X_2\right] \\ \operatorname{cov}\left[X_2, X_1\right] & \operatorname{cov}\left[X_2, X_2\right] \end{pmatrix} \stackrel{\text{(a)}}{=} \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix}$$

¹ почему?

называется ковариационной матрицей вектора Х. ②

Пусть $Q(z_1, z_2) \stackrel{\text{def}}{=} \mathbf{z}' \mathbf{\Lambda} \mathbf{z}, \ \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$. Такое выражение называется $\kappa \epsilon a \partial p a - m u u h o \tilde{u}$ от \mathbf{z} . Квадратичную форму можно записать в явном виде:

$$Q(z_1, z_2) = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}' \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 z_1 + z_2 \rho \\ \rho z_1 + \sigma_2^2 z_2 \end{pmatrix}' \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$
$$= \sigma_1^2 z_1^2 + 2\rho z_1 z_2 + \sigma_2^2 z_2^2.$$

Утверждение 1. Квадратичная форма Q является неотрицательно определенной, то есть $Q(z_1, z_2) \geq 0$ для любых z_1 и z_2 .

Замечание 1. Матрица Λ , для которой $Q(\mathbf{z})>0$ для всех \mathbf{z} называется положительно определенной.

Доказательство. Мы рассмотрим случай $\sigma_2^2 \neq 0$. Утверждение очевидно, если $z_1=0$ и $z_2=0$. Если же, например, $z_1\neq 0$, то

$$Q(z_1, z_2) = z_1^2 (\sigma_1^2 + 2\rho t + \sigma_2^2 t^2), \qquad t = z_2/z_1.$$

Квадратный трехчлен в скобках представим в виде

(1)
$$\sigma_1^2 + 2\rho t + \sigma_2^2 t^2 = \left(\sigma_2 t + \frac{\rho}{\sigma_2}\right)^2 + \sigma_1^2 - \frac{\rho^2}{\sigma_2^2}.$$

Теперь видно, что $Q(z_1,z_2)\geq 0$, так как $\rho^2\leq \sigma_1^2\sigma_2^2$ в силу неравенства Коши-Буняковского. \square

1.4. Неравенство Коши-Буняковского. Последний шаг в доказательстве утверждения 1 опирается на неравенство

(2)
$$(\mathsf{E}\left[\xi\eta\right])^2 \le \mathsf{E}\left[\xi^2\right] \cdot \mathsf{E}\left[\eta^2\right],$$

которое верно для любых случайных величин ξ и η с конечными вторыми моментами.

Для доказательства неравенства (2) заметим, что $(s\xi+\eta)^2\geq 0$ для любого s. Поэтому $\mathsf{E}\left[(s\xi+\eta)^2\right]\geq 0$. Значит $\mathsf{E}\left[\xi^2\right]s^2+2\mathsf{E}\left[\xi\eta\right]s+\mathsf{E}\left[\eta^2\right]\geq 0$ для любого s. Следовательно, дискриминант неположителен: $(\mathsf{E}\left[\xi\eta\right])^2-\mathsf{E}\left[\xi^2\right]\mathsf{E}\left[\eta^2\right]\leq 0$, что и требовалось доказать.

 ${f 3}$ адача * 1. Как доказать утверждение 1 при $\sigma_2^2=0$?

Задача* 2. Может ли выполняться равенство $Q(z_1, z_2) = 0$ и для каких z?

 $^{^2}$ объяснить (a)

2. Многомерный гауссовский случайный вектор

Определение 2. Случайный вектор **X** размерности n называется n-мерным гауссовским случайным вектором, если случайная величина $c_1X_1 + \cdots + c_nX_n$ является гауссовской при любых $c_1, \ldots, c_n \colon |c_1| + \cdots + |c_n| \neq 0$.

Как и для двумерных векторов справедливы такие свойства:

- (G_1) гауссовские векторы существуют для любой размерности n: например, вектор \mathbf{X} , составленный из независимых стандартных гауссовских случайных величин X_1, \ldots, X_n , является гауссовским;
- (G_2) любая из случайных величин X_1, \ldots, X_n является гауссовской, если \mathbf{X} гауссовский случайный вектор;
- (G_3) если **X** гауссовский случайный вектор, то $X_1 + \cdots + X_n$ является гауссовской случайной величиной;
- (G_4) любое маргинальное распределение вектора ${f X}$ является гауссовским.

Напомним, что маргинальным распределением случайного вектора $\mathbf{X} = (X_1, \dots, X_n)'$ называется распределение его подвектора $(X_{i_1}, \dots, X_{i_m})'$, m < n.

Задача 3. Доказать, что случайный вектор размерности n имеет 2^n-2 маргинальных распределений.

Задача 4. Доказать свойства (G_1) - (G_3) .

Доказательство свойства (G_4) . Ограничимся случаем маргинального распределения случайного подвектора, составленного из первых m < n координат вектора \mathbf{X} . Докажем, что оно является m-мерным гауссовским. Действительно, пусть c_1,\ldots,c_m — произвольные константы, $|c_1|+\cdots+|c_m|\neq 0$. Тогда $c_1X_1+\cdots+c_mX_m$ является гауссовской случайной величиной, так как $c_1X_1+\cdots+c_mX_m=c_1X_1+\cdots+c_mX_m+0\cdot X_{m+1}+\cdots+0\cdot X_n$.

2.1. Векторные обозначения. Векторная нотация более экономна при изучении гауссовских векторов. Например, $c_1X_1+\dots+c_nX_n=\mathbf{c}'\mathbf{X}$. В дальнейшем мы используем векторные обозначения и матричные операции. Векторы понимаются как вектор-столбцы. Если матрица имеет m строк и n столбцов, то говорим о $m\times n$ -матрице. Транспонирование матрицы \mathbf{B} обозначаем \mathbf{B}' . В частности, если \mathbf{X} — это n-вектор (то есть $n\times 1$ -матрица ③), то $\mathbf{X}\mathbf{X}'$ — это $n\times n$ -матрица. ④

Если все элементы b_{ij} матрицы ${\bf B}$ являются случайными величинами, то мы говорим о случайной матрице. Если b_{ij} — это элементы матрицы ${\bf B}$ и все они имеют математические ожидания, то через ${\sf E}[{\bf B}]$ мы обозначаем матрицу, составленную из ${\sf E}[b_{ij}]$. В частности, если ${\bf X}$ — это случайный n-вектор (случайная $n \times 1$ -матрица), то ${\sf E}[{\bf X}]$ — это вектор составленный из ${\sf E}[X_i]$.

Пемма 1. Пусть X — это случайный n-вектор, a b — неслучайный n-вектор. Тогда E[X+b]=E[X]+b.

Доказатель ство. Вытекает из линейности математического ожидания случайных величин. ⑤ □

 $^{^3}$ почему не $1 \times n$ матрица?

⁴проверить!

⁵доказать!

Лемма 2. Пусть A — случайная матрица, а B и C — неслучайные матрицы, согласованных с A размерностей. Тогда $E[BA] = B \cdot E[A]$ и $E[AC] = E[A] \cdot C$. В частности, если X — случайный вектор, B — неслучайная матрица и $\mu = E[X]$, то $E[BX] = B\mu$.

Доказательство. Обозначим элементы (i,j) матриц $\mathbf{B}\mathbf{A}$ и $\mathbf{B}\cdot\mathsf{E}[\mathbf{A}]$ через ξ_{ij} и e_{ij} соответственно. Тогда в силу линейности математического ожидания для случайных величин

$$\xi_{ij} = \sum_{\nu} b_{i\nu} a_{\nu j}, \quad e_{ij} = \sum_{\nu} b_{i\nu} \mathsf{E}\left[a_{\nu j}\right], \quad \text{откуда} \quad \mathsf{E}\left[\xi_{ij}\right] = e_{ij}.$$

Это и доказывает первое утверждение леммы 2.

Задача 5. Доказать второе утверждение леммы 2.

2.2. Ковариационная матрица. Пусть \mathbf{X} — случайный n-вектор. Предположим, что существуют вторые моменты у всех случайных величин X_i (то есть у всех координат вектора \mathbf{X}).

Определение 3. Матрица Cov[X], составленная из ковариаций $cov[X_i, X_j]$, $1 \le i, j \le n$, называется ковариационной матрицей вектора X.

Лемма 3. Пусть X — случайный n-вектор. Обозначим $E[X] = \mu$ и пусть Λ — ковариационная матрица вектора X. Тогда $\Lambda = E[(X - \mu)(X - \mu)']$. В частности, Λ является симметричной матрицей.

Напомним, что свойство симметричности матрицы $\mathbf{\Lambda} = \left(\lambda_{ij}\right)_{i,j=1}^n$ означает, что $\lambda_{ij} = \lambda_{ji}$.

Доказательство. Поскольку

$$(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})' = \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \dots \\ X_n - \mu_n \end{pmatrix} (X_1 - \mu_1, X_2 - \mu_2, \dots, X_n - \mu_n)$$

$$= \begin{pmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \dots & (X_1 - \mu_1)(X_n - \mu_n) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \dots & (X_2 - \mu_2)(X_n - \mu_n) \\ \dots & \dots & \dots & \dots \\ (X_n - \mu_n)(X_1 - \mu_1) & (X_n - \mu_n)(X_2 - \mu_2) & \dots & (X_n - \mu_n)^2 \end{pmatrix},$$

то $\mathsf{Cov}[\mathbf{X}] = \mathsf{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})']$. Симметричность ковариационной матрицы вытекает из ее определения. \square

Теорема 1. Ковариационная матрица является неотрицательно определенной, то есть $Q(\mathbf{z}) \stackrel{\text{def}}{=} \mathbf{z}' \Lambda \mathbf{z} \geq 0$ для любого неслучайного вектора \mathbf{z} .

Доказательство. Из лемм 2 и 3 вытекает, что

$$\mathbf{z}' \mathbf{\Lambda} \mathbf{z} = \mathbf{z}' \mathsf{E} \left[(\mathbf{X} - \boldsymbol{\mu}) (\mathbf{X} - \boldsymbol{\mu})' \right] \mathbf{z} = \mathsf{E} \left[\mathbf{z}' (\mathbf{X} - \boldsymbol{\mu}) (\mathbf{X} - \boldsymbol{\mu})' \mathbf{z} \right].$$

Для любых векторов **a** и **b** выполнено $(\mathbf{a}'\mathbf{b})' = \mathbf{b}'\mathbf{a}$. © Поэтому

$$\mathbf{z}' \mathbf{\Lambda} \mathbf{z} = \mathsf{E} \left[\boldsymbol{\xi}' \boldsymbol{\xi} \right], \qquad \boldsymbol{\xi} \stackrel{\mathrm{def}}{=} (\mathbf{X} - \boldsymbol{\mu})' \mathbf{z}.$$

 $^{^6}$ проверить!

Поскольку $\boldsymbol{\xi}'\boldsymbol{\xi}=\xi_1^2+\cdots+\xi_n^2$, то $\mathbf{z}'\mathbf{\Lambda}\mathbf{z}\geq 0$, что и требовалось доказать. \square

Замечание 2. Утверждение 1 (частный случай теоремы 1 для n=2) было доказано с помощью неравенства Коши–Буняковского (см. §1.4). В свою очередь, неравенство Коши–Буняковского вытекает из теоремы 1. Действительно, пусть n=2. Тогда, как доказано в теореме 1, $Q(z_1,z_2)\geq 0$ для любых z_1,z_2 . Квадратичная форма $Q(z_1,z_2)$ имеет вид (1). Подставляя $t=-\frac{\rho}{\sigma_2^2}$ в (1), видим, что $\sigma_1^2-\frac{\rho^2}{\sigma_2^2}\geq 0$. А это и есть неравенство Коши–Буняковского (2).

Следствие 1. Пусть X — случайный n-вектор, b — неслучайный n-вектор. Тогда Cov[X+b]=Cov[X].

Доказательство. Обозначим $\mathbf{Y} \stackrel{\text{def}}{=} \mathbf{X} + \mathbf{b}$. Тогда матрица $\text{Cov}\left[\mathbf{X} + \mathbf{b}\right]$ составлена из чисел $\text{cov}\left[Y_i, Y_j\right]$, а матрица $\text{Cov}\left[\mathbf{X}\right]$ — из $\text{cov}\left[X_i, X_j\right]$. Так как $\text{cov}\left[Y_i, Y_j\right] = \text{cov}\left[X_i, X_j\right]$, 0 то следствие доказано. \square

Лемма 4. Пусть X — случайный n-вектор, B — неслучайная $m \times n$ -матрица. Обозначим $\mu = \mathsf{E}[X]$, $\Lambda = \mathsf{Cov}[X]$. Тогда $Y = \mathbf{B}X$ — случайный твектор, для которого $\mathsf{E}[Y] = \mathbf{B}\mu$, $\mathsf{Cov}[Y] = \mathbf{B}\Lambda\mathbf{B}'$.

Доказательство. Утверждение о математическом ожидании вытекает из леммы 2. Она же вместе с леммой 3 позволяет доказать и утверждение о ковариационной матрице:

$$\mathsf{Cov}\left[\mathbf{Y}\right] = \mathsf{E}\left[\left(\mathbf{Y} - \mathsf{E}\left[\mathbf{Y}\right]\right)\left(\mathbf{Y} - \mathsf{E}\left[\mathbf{Y}\right]\right)'\right] = \mathsf{E}\left[\left(\mathbf{B}\mathbf{X} - \mathbf{B}\boldsymbol{\mu}\right)\left(\mathbf{B}\mathbf{X} - \mathbf{B}\boldsymbol{\mu}\right)'\right]$$

$$\stackrel{(c_1)}{=} \mathsf{E}\left[\left(\mathbf{B}(\mathbf{X} - \boldsymbol{\mu})\right)\left(\mathbf{X} - \boldsymbol{\mu}\right)'\mathbf{B}'\right] \stackrel{(c_2)}{=} \mathbf{B}\boldsymbol{\Lambda}\mathbf{B}'.$$

8

2.3. Линейные преобразования гауссовских векторов. Если гауссовский вектор **X** имеет вектор средних μ и матрицу ковариаций Λ , то мы пишем $\mathbf{X} \in \mathcal{N}(\mu, \Lambda)$. Иногда мы также пишем $\mathbf{X} \in \mathcal{N}_n(\mu, \Lambda)$, чтобы отметить размерность векторов **X**, μ и матрицы Λ .

Заметим, что математические ожидания и ковариации координат гауссовского вектора существуют в силу свойства (G_2) .

Теорема 2. Пусть **B** матрица размерности $m \times n$, **b** — вектор размерности m. Если случайный n-вектор **X** является гауссовским, причем $\mathbf{X} \in \mathbb{N}_n (\mu, \mathbf{\Lambda})$, то случайный m-вектор $\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{b}$ также является гауссовским, причем $\mathbf{Y} \in \mathbb{N}_m (\mathbf{B}\mu + \mathbf{b}, \mathbf{B}\mathbf{\Lambda}\mathbf{B}')$.

9

Доказательство. Сначала докажем, что \mathbf{Y} является гауссовским вектором. Для любого вектора \mathbf{c}' имеем $\mathbf{c}'\mathbf{Y} = \mathbf{c}'\mathbf{B}\mathbf{X} + \mathbf{c}'\mathbf{b} = \mathbf{u}'\mathbf{X} + v$, где $\mathbf{u} = \mathbf{B}'\mathbf{c}$ и $v = \mathbf{c}'\mathbf{b}$. $\mathbf{w} = \mathbf{c}'\mathbf{b}$. Случайная величина $\mathbf{u}'\mathbf{X}$ является гауссовской по определению 2, а v — действительной константой. Поэтому и $\mathbf{c}'\mathbf{Y}$ является гауссовской случайной величиной. Утверждение о математическом ожидании вытекает из лемм 1

⁷доказать!

⁸почему выполнены (c_1) и (c_2) ?

 $^{^9 \}Pi$ очему вектор **Y** имеет размерность m?

¹⁰почему?

и 2. Следствие 1 и лемма 4 доказывают утверждение о ковариационной матрице. $\ \square$

 $^{^{\}dagger}$ Всего в тексте было 10 вопросов